The Use of Census Journey-to-Work Data In Ontario Municipalities

B.G. Hutchinson, Professor
D.P. Smith, Research Assistant
Department of Civil Engineering
University of Waterloo

Published by
The Research and Development Branch
Ontario Ministry of Transportation and
Communications
Hon. James W. Snow, Minister
H.F. Gilbert, Deputy Minister

This study was commissioned under the Ontario Joint
Transportation and Communications Research
Programme (O.J.T. & C.R.P.)

For additional copies write:

The Editor
Research and Development Branch
Ministry of Transportation and Communications
1201 Wilson Avenue
Downsview, Ontario
M3M 1J8

June 1981

Abstract

This report integrates the findings of two research projects that analyzed the 1971 Census of Canada Journey-to-Work data. Undertaken at the University of Waterloo and sponsored by the Ontario Ministry of Transportation and Communications, the first project was completed in January 1978; the second, in February 1979. Their primary objective was to explore the extent to which the census Journey-to-Work data might be used routinely for transport planning studies in Ontario census areas. A second objective was to use this information, which had been systematically collected across all census areas, to improve our understanding of urban spatial structures and the journey to work. The earlier reports described in detail the research process and findings. The aim of this report is to provide an overview of these findings in terms that might be more useful to transport planners.

The Journey-to-Work data from the 1971 Census of Canada provided a major breakthrough for urban transport planning, since it was the first time that consistent work trip travel information was available for all of the Ontario census areas simultaneously. The value of the 1971 Census journey-to-work information will be increased dramatically if comparable data are collected in the 1981 Census of Canada. This would provide transport planners with a longitudinal view of the growth and change in the spatial patterns of work trip travel demands over a decade - information that has never been available in Canadian urban areas. Transport policies and investments must be evaluated over long time periods and a satisfactory view of the longitudinal character of urban travel demands has never been available. Future travel demands are estimated by computer-based models, but the capabilities of these models have never been tested.

Contents

Executive Summary

This report describes a series of analyses of the journey-to-work data obtained in the 1971 Census of Canada. The 1971 Census was the first Canada census in which the type and location of employment of the residential labour force were determined. Census household data are already widely used in urban studies, and the availability of information on the home to work linkages of urban residents provided a very important extension to the household data. The 1971 Census provided the first systematic and comprehensive set of data on the spatial distributions of employment by industry sector for Canadian urban areas. The unavailability of good employment data has been a major deficiency in virtually all the urban transport studies conducted in Canada.

The 16 Ontario census areas (CAs) range in size from the Guelph census agglomeration to the Toronto census metropolitan area (CMA). The basic spatial unit to which the 1971 journey-to-work data are coded in these census areas is the census tract where the average census tract contains about 5 000 people. In 1971 Guelph was subdivided into 14 official census tracts; there were 452 in the Toronto CMA. The contiguous urbanized portions of a census area are covered by the official census tract system. The analyses described in this report deal with the properties of and the interactions between official census tracts. Home to work linkages occur between official census tracts and the surrounding census districts, and with other census areas. But in most census areas, the linkages between official census tracts account for about 80% of the total linkages. The coded journey-to-work data available on a routine basis from Statistics Canada represent approximately an 11% sample of the home to work linkages made by the households in each census area. The term "linkages" is used rather than trips, since some members of the labour force may make more than one trip per day between home and work.

The census journey-to-work data have been supplemented by information on the properties of the road networks of each census area. Road networks have been coded at roughly the arterial/subarterial level and the length of each road link measured. Road network distances have been used in most of the analyses, since consistent travel speed data were not available for all census areas.

The Ontario census areas vary widely in population size, economic base, spatial character and geographic setting. In 1971, six census areas had populations of less than 100 000; three had populations between 100 000 and 200 000; four had populations between 200 000 and 300 000; and three had populations of 500 000 or greater. In 1971, three broad industry groups accounted for about 70% of the jobs: and these were manufacturing, trade and services. The principal variations in the economic bases between the Ontario census areas were in primary (mining), manufacturing, transportation and the public administration sectors. In Sudbury 26.5% of the jobs were in mining; Brantford, Oshawa and Kitchener had strong manufacturing industry bases. Employment in transportation was well above the provincial average in Thunder Bay, while 16.3% of the labour force in Ottawa was in public administration.

The spatial distributions of employment in the Ontario census areas reflect these variations in employment base. Employment in the census areas with light manufacturing, such as Brantford and Kitchener, tends to disperse throughout the area. Census areas specializing in office based employment, such as Ottawa, tend to have strong concentrations of CBD employment. Eight of the Ontario census areas have heavy concentrations in two or three census tracts. In four of these areas, residential census tracts are located away from the job centres, creating longer average trip lengths. The spatial arrangement of houses and jobs are influenced significantly by topographic constraints in nine of the census areas. The most common constraint is being located alongside a water body. In several census areas these effects work in combination to produce long average trip lengths unrelated simply to population size.

The per capita amount of home to work travel in the Ontario census areas varies from 1.24 km per capita in the Peterborough CA to 3.93 km per capita in the Toronto CMA. Although the census areas tend to cluster around a trend line showing increasing per capita travel with increasing population, there are some significant deviations from the trend. For example, the per capita amount of travel in Sarnia is about 32% higher than the mean for the six census areas in its size group, and the amount of travel in Sudbury is about 27% higher than the mean for the three census areas in its size group. In contrast, the amount of travel in those census areas with widely distributed concentrations of manufacturing employment is lower than average.

The variations in the per capita home to work travel distance are also reflected in variations in the mean home to work distance, which varies from 3.6 km in Peterborough to 10.8 km in the Toronto CMA. The total amount of home to work travel is also influenced by the community's activity rate, which varies from 37.6% of the population in Oshawa to 45.0% in Sudbury and Toronto.

The differences in the total amount of home to work travel between urban areas in the same size class reflect the variations in the spatial structures of communities. Employment opportunities dispersed throughout an area along with appropriate housing opportunities contribute to the development of localized commuting patterns, particularly where the employment opportunities create few unsatisfactory environmental impacts. This effect may be observed in contour maps of the mean trip lengths from residential census tracts. Guelph, Peterborough, Brant-

ford, Oshawa, Kitchener, London and St. Catharines exhibit this balance between the spatial distributions of houses and jobs. In contrast, Sarnia, Sault Ste. Marie, Kingston, Sudbury, Windsor and Hamilton have strong linkages between employment concentrations and residential areas concentrated in one or two spatially separated sectors.

This comparative analysis of the Ontario census area demonstrates objectively the opportunities that are available for reducing travel through enlightened urban spatial management.

Work trip travel patterns may be synthesized from land use data if the spatial distributions of labour force and employment by census tract can be estimated first. Spatial variations in the amount of labour force in each census tract have been analyzed in terms of population, households and dwelling unit composition. The most reliable regression equations obtained were in terms of the dwelling unit composition of each census tract. Regression equations developed in terms of the number of single detached dwelling units and the number of attached dwelling units explained from 90 to 99% of the observed census tract to census tract variation in labour force. The partial regression coefficients of the number of attached dwelling units varied from 1.50 to 1.57 labour force per dwelling unit except for Thunder Bay, Sudbury, Kitchener and Ottawa, where the activity rates varied widely from the provincial average. The partial regression coefficients of the number of attached dwelling units exhibited more variability, ranging from 0.83 labour force per dwelling unit in the St. Catharines CMA to 1.40 in the Toronto CMA.

Some of the variability in the partial regression coefficient magnitudes is statistical in nature due to the lower number of multiple dwelling units in many communities. However, some of the variability is due no doubt to differences in the housing market between census areas. The explanatory power and consistency of these regression equations across the Ontario census areas suggests that work trip production studies should be standardized in terms of dwelling unit composition. Dwelling unit composition may be reliably estimated for future time horizons from land use planning policies.

Special tabulations of home to work linkages stratified by tenancy status, period of residence and car ownership have allowed stratified census tract labour force prediction equations to be developed. Of particular interest are the prediction equations developed for estimating transit captive and noncaptive labour force rates. Statistically significant prediction equations were developed for each census area, but the partial regression coefficients of the two dwelling unit types do vary between census areas. Average transit captive generation rates from single attached dwelling units are from two to three times greater than from single detached dwelling units. The noncaptive rate is about three times larger from detached dwelling units than from attached dwelling units.

Trip attraction equations cannot be established from the census data, since independent measures of land use at employment locations are not available from the census. However, opportunities exist for developing appropriate relationships in those census areas with reliable land use data coded to the census tract level.

The cell entries in observed home to work linkages matrices reflect two types of effects: (1) the rate of interaction between a pair of zones; and (2) the absolute magnitudes of labour force and employment in the two interacting zones. The basic hypothesis of the gravity model is that the rate of interaction between zones is a function of the spatial separation of zones. To better understand the spatial interaction patterns in the Ontario census areas, it is important to extract from the home to work linkages matrices the pure spatial interaction effects that are independent of zone size effects. This better understanding would allow improved forms of the gravity model to be developed.

Bi-proportional balancing techniques were used to transform the home to work linkages matrix of each census area so that the entries in each row and in each column sum to a constant magnitude. The effect of this operation is to produce a linkages matrix in which the entries represent the interaction magnitudes that would occur between zones of equal size, or the pure interaction effects between home and work zones. This bi-proportional balancing process is straightforward, i.e., the observed matrix row entries are each scaled by the same proportion so that each row total is equal to an arbitrary constant. The columns of this adjusted matrix are then summed and scaled so that each column total is equal to the arbitrary constant. This row by column adjustment process is repeated until convergence is obtained with each row and column total being equal to the arbitrary constant. If the arbitrary constant were equal to one, the cell entries of the adjusted matrix could be interpreted as the unit probabilities of interaction between zones.

The inter-zonal interaction effects isolated in this way were examined to isolate patterns of interaction. A clustering technique was used to group residential census tracts on the basis of similarities in their destination vectors in the bi-proportionally adjusted matrix. The procedure begins with 'n' census tracts and groups them into a sequence of clusters until all census tracts are merged into a single cluster. The heterogeneity of the census tracts within each cluster is represented by the within cluster sum of squared differences in the destination vectors of the census tracts. The sum of this quantity over all clusters indicates the overall heterogeneity in the commuting

patterns at each stage in the clustering process. The error sum of squares increased slowly during the first stages of the clustering process and then increased sharply when residential zones with quite different destination vectors were forced into the same cluster. The hierarchical structure of census tract merging in each census area is represented in a dendrogram. The dendrograms show the sequence in which zones cluster and the error sums of squares magnitudes at which they cluster.

Commuting subregions were identified for the census areas from these cluster analyses. There are five broad determinants of these subregions:
(i) multi-community composition of a census area;
(ii) topographic and man-made constraints;
(iii) the time sequence of urban development;
(iv) socio-economic factors; and
(v) the domination of large employment centres.

The Thunder Bay, Kitchener, St. Catharines and Hamilton census areas provide examples of the impact that municipal boundaries have on commutersheds in almost all cases of clustering. At the 4 or 5 cluster level the cluster boundaries were coincident with municipal boundaries. The London, Hamilton and Ottawa census areas provide examples of the impacts that topographic features have on commuting patterns. Timing of development influences were detected in almost all the census areas, particularly those that grew rapidly during 1966–1971. The socio-economic characteristics of the households are the most common determinant of commuting patterns across all census areas. In many cases, the other determinants mentioned above co-vary with socio-economic influences.

A prerequisite of any systematic study of the gravity model is the development of a suitable goodness of fit statistic for comparing the cell entries of the observed and simulated trip matrices. The coefficient of determination is the most commonly used goodness of fit statistic used in transport systems analysis. Previous studies have shown that this statistic is completely unsuitable for use for this problem. The effectiveness of a variety of goodness of fit statistics was explored by a series of simulation experiments for eight of the larger Ontario census areas. Simulated trip matrices were calculated from the observed matrices by multiplying each of the observed cell entries by a randomly generated error with a specified range. These analyses showed that the best goodness of fit statistic to use is the phi-statistic of information theory. The phi-statistic increases with increasing numbers of linkages at a constant percent error, while the phi-statistic divided by the number of linkages was roughly constant across all census area sizes at the same error magnitude.

A variety of modifications to the gravity model was tested to improve the goodness of fit for the Ontario census areas. A very flexible gravity model computer program was developed that allows a wide range of gravity structures to be tested and exhaustive goodness of fit statistics to be calculated. In most cases, gravity models were estimated only for the eight largest Ontario census areas other than Toronto.

Production-constrained, attraction-constrained and doubly-constrained versions of the gravity model were calibrated using road network distances in kilometres as the measure of generalized travel costs between zones. This basic set of models was calibrated to provide the datum against which modifications to the gravity model designed to capture some of the other determinants of commuting patterns might be assessed. The calibration criterion used for all model types was the minimization of the sum of the absolute differences in the ordinates of the observed and simulated trip length frequency distributions. In most cases, this criterion also satisfied the normally used criterion of equality of observed and simulated area-wide mean trip lengths.

With the normalized phi-statistic as the goodness of fit measure, the calibrated models produced trip tables of the same statistical quality as those produced by the random error generation with an error range of 75 to 100%. A comparison of the model types showed that the doubly-constrained model was superior to the two singly-constrained models for all census areas, with the attraction-constrained model exhibiting the poorest goodness of fit. The goodness of fit deteriorated with increasing urban area population. The deterrence function parameter tended to decrease with increasing population size, but there is a significant amount of variability associated with this trend. Although the doubly-constrained model provided the best goodness of fit, there are real difficulties in providing any reasonable interpretation of the parameter and balancing factor magnitudes.

Doubly-constrained models using network travel times rather than network distances were calibrated for the Kitchener, Hamilton and Ottawa census areas and compared with those developed using network distances. The phi-statistics showed that the models using travel time were marginally inferior to those using network distances. Analyses of the origin zone-specific goodness of fit statistics showed that in the Ottawa CMA the time-based models improved the goodness of fit of some of the outlying residential census tracts, but that this was not a general trend across all census tracts.

Inter-zonal generalized travel costs were also created from various combinations of network distances and network times for the Kitchener CMA. Analyses of the goodness of fit statistics showed that doubly-constrained models calibrated using these generalized cost functions are superior to those using travel times, but marginally inferior to the models calibrated using network distances.

Detailed comparisons of the trip interchange residuals for the Ottawa CMA and a partitioning of the phi-statistic into over and underestimation residuals by intra-zonal and inter-zonal trip interchanges showed that the major source of error is the underestimation of trip interchange magnitudes and that the superior behaviour of the doubly-constrained model is due to a reduction in the underestimation residuals. However, this is not a completely unbiased interpretation of the residuals, since earlier simulation studies had shown that the phi-statistic is more sensitive to underestimation errors than to overestimation errors.

A number of stratified gravity models were calibrated using all three basic forms of the gravity model. The three groups of stratified gravity models calibrated were:

(i) multi-parameter gravity models with separate travel deterrence function parameters estimated for each census tract;

(ii) subregion specific models in which separate travel deterrence parameter magnitudes are calibrated for from four to eight commuting subregions depending on the specific area; and

(iii) a set of models calibrated for separate socio-economic groups.

The multi parameter production-constrained models all exhibited superior goodness of fit characteristics to the single parameter production-constrained models, but were not as good as the single parameter doubly-constrained models. The largest improvements occurred in the multi-community census areas of Kitchener and St. Catharines. The multi-parameter attraction-constrained models were inferior to the single parameter versions except for Thunder Bay, Kitchener and St. Catharines. The multi-parameter doubly-constrained models all exhibited superior goodness of fit characteristics to the single parameter versions.

The cluster analyses of the census area formed the basis for the identification of the calibration subregions. These ranged from three in Thunder Bay and Kitchener to eight in the Ottawa census area. The subregion specific production-constrained models are superior to the single parameter versions, but are inferior to both the single and multi-parameter production-constrained models. An analysis of the subregion specific doubly-constrained models showed that they performed marginally better than the single parameter version, but are inferior to the multi-parameter models.

The most useful models stratified by socio-economic characteristics are those developed for non-car owners and car owners. Detailed studies for the Kitchener area showed that doubly-constrained versions of these stratified models perform at a superior level to the single parameter doubly-constrained models.

These studies of the gravity model have shown that existing models do not provide reasonable estimates of observed base year trip tables. It seems clear that horizon year travel demands derived from the gravity model would have even larger sources of error. The errors are so large that there seems little point in continuing to use the gravity model in its present form in urban transport planning studies.

1/ Introduction

The practical importance of the journey-to-work data collected in the 1971 Census of Canada has long been recognized by the Ontario Ministry of Transportation and Communications. It has used census information in a number of transport model building and research projects conducted both within the Ministry and by external agencies. This report describes the results of a research program, sponsored at the University of Waterloo by the Ontario Ministry of Transportation and Communications, which was concerned with a detailed analysis of the spatial interaction patterns observed in the 1971 Census of Canada for all the Ontario census areas other than the Toronto CMA.

A principal advantage of including the "Journey-to-Work" question in the decennial household census is that travel information may be collected far more cheaply and reliably during the census than by special purpose surveys conducted from time to time within individual communities. Census population and household data coded to the census tract level are already widely used in urban studies, and the inclusion of place of work information would provide an important extension to the household data. Information would not only be available on the spatial linkages between home and work, but also on the spatial distribution of employment by industry sector. The unavailability of information on the spatial distributions of employment by industry sector is perhaps the most important deficiency of current urban data bases.

The availability of the journey-to-work information for the same year for all Canadian census areas provides a unique opportunity for the comparative analysis and transfer of information between urban areas. A further advantage of regularly collected information on the journey to work is that it would provide a time series view of urban spatial structure and work trip demands. Urban areas are dynamic systems, and sound urban transport policy must be developed from a clear understanding of urban growth and change.

1.1/ Research Program Objectives

Two broad goals were set for this research program:

1/To develop a better understanding of work trip spatial interaction patterns across a wide range of urban areas and the factors that influence these interaction patterns.

2/To explore the use of the census journey to work data in the development of generalized transport models that might be used to synthesize future work trip demands without the collection of data in comprehensive special purpose surveys.

The specific objectives set within these broad goals were as follows:

1/To calculate several aggregate indicators of the amount of work trip travel in Ontario census areas and to interpret any differences in these indicators in terms of the scale and spatial characteristics of the areas.

2/To develop regression equations for estimating the rate of generation of work trip linkages from residential areas and to explore the extent to which these equations might be generalized across all Ontario census areas.

3/To develop regression equations for estimating the rate of generation of work trip linkages stratified by socio-economic group from residential areas.

4/To use bi-proportional matrix balancing techniques to calculate normalized trip linkage magnitudes between each pair of origin-destination zones in order to separate interaction effects from zone size effects.

5/To use clustering techniques to identify groups of census tracts with similar interaction behaviour in order that factors influencing spatial interaction may be isolated and incorporated in the gravity model structure.

6/To develop a more flexible gravity model computer program that allows alternative types of gravity model to be calibrated and model errors to be examined in detail.

7/To examine alternative statistical criteria that may be used to assess the goodness of fit of trip distribution models.

8/To estimate the parameter magnitudes and goodness of fit of alternative forms of the gravity model that may be stratified by geographic subregion, socio-economic group, etc.

Complete analyses are reported for the 15 Ontario census areas other than the Toronto CMA, which was excluded from these analyses because of computer costs. However, some results are included for the Toronto CMA.

2/ The Role of Models in Urban Transport Planning

The initial round of urban transport planning studies performed in Ontario municipalities during the late 1950s and early 1960s followed a rather tightly organized format. The process usually began with inventories of existing travel facilities, land use and travel demands, followed by the development of a transport systems analysis model that was used to estimate future travel demands. These forecasts of future travel demands were then applied to develop alternative transport networks whose probable future operating characteristics were analyzed again using the calibrated transport demand model. Capacity deficiencies were then isolated and integrated into a 10-to-20-year investment program.

During the 1970s the structure of the urban transport planning process in many Ontario municipalities evolved from the earlier process. Perception of the urban transport problem varied between municipalities, and different study structures were developed in response to the different problem perceptions. Broad strategic type issues were emphasized in some municipalities; in others, shorter run traffic engineering and transit operating problems were addressed. In some municipalities, recent transport planning studies have changed very little from the earlier studies. (The recently completed study in the Regional Municipality of Waterloo provides one example [1].)

2.1/ Typical Model Structures

The broad sequence of steps involved in the typical urban transport planning study is illustrated in a generalized way in Figure 1, where the role played by calibrated transport demand models is emphasized. The demand models are used to estimate future travel demands as well as the network equilibrium demands likely to result from the interaction between demand and a particular supply strategy.

Urban transport systems models attempt to capture parts of the location decision-making behaviour of urban residents as well as their transport decision-making behaviour and the ways in which this behaviour might be influenced by changes in the transport policy environment affecting individual trip makers. Trip-making behaviour is usually represented by a sequence of submodels of the following form:

$$T_{ij}^{mr} = O_i \cdot S_{ij} \cdot S_{ij}^{m} \cdot S_{ij}^{mr} \qquad (1)$$

where:

T_{ij}^{mr} = the number of trips from origin zone i into destination zone j by mode m and network route r

O_i = the number of trips produced in zone i

S_{ij} = the proportion of trips produced in zone i that travel to destinations in zone j

S_{ij}^{m} = the proportion of trips between zones i and j that travel by mode m

S_{ij}^{mr} = the proportion of trips between zones i and j that travel by mode m and route r

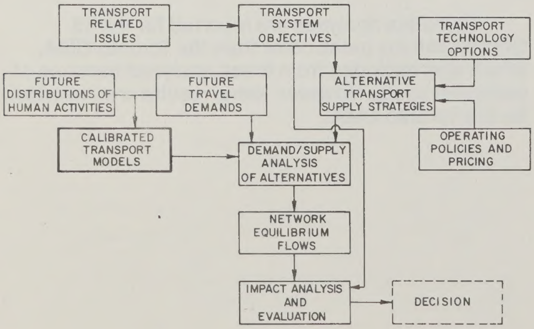
The travel demand matrices calculated by Equation 1 are conditional, i.e.,they are for a particular land use plan and transport network proposal. Equation 1 is typically applied for a number of trip purposes including work, shopping, social recreational, etc.

The first submodel in Equation 1 is trip generation, which has largely been a matter of empirical investigation in which observed rates of trip generation have been related to a variety of land use variables by regression analysis. The trip distribution submodel, S_{ij}, estimates the *share* of the trips generated at each origin zone likely to go to each of the possible destination zones. The gravity model is typically used to estimate these shares. The character of the modal split submodel, S_{ij}^{m}, has varied widely between studies. In studies performed in most medium and smaller Ontario municipalities, no serious attempts have been made to estimate modal split with constant proportions of trips by transit for all zones being assumed. In the larger cities, some form of diversion curve has usually been developed that attempts to capture the cost and service characteristics of the alternative modes. Finally, the route assignment submodel, S_{ij}^{mr}, has usually been of the all-or-nothing type where all of the road travel demand between a particular origin-destination pair is assigned to the minimum travel time path between the origin-destination pair.

Much has been written about the deficiencies of this four-stage transport demand model from both the theoretical and empirical points of view. These criticisms will not be reviewed in this report, but it is important to highlight some of the difficulties and inconsistencies experienced in applying this model set to Ontario municipalities.

Perhaps the overriding problem is the limited data base typically available for calibrating the model.

Figure 1/ Broad Structure of Transport Planning Process



In many of the recent studies, data bases have either been synthesized from earlier travel surveys or obtained from very small samples of travel (1 to 2% of households). Data bases of this type provide some information for estimating model parameters, but they certainly do not provide adequate information for assessing the goodness of fit of models. As a result, most studies rely on model estimated and observed screenline volume comparisons for the evaluation of the integrity of the models. Any major differences revealed by this comparison are typically removed by arbitrary changes to the network distances and times used in the distribution and assignment models. Independent checks at each stage of the modelling process are not possible because of the absence of a good data base.

The use of models calibrated in this way to estimate future travel demands is clearly untenable. It is unlikely that the changes made in the model to force it to reproduce base year flows will also apply to the horizon year when the spatial structure of an urban area will be different. The goodness of fit qualities of the traditional transport demand models have never been adequately assessed for Ontario urban areas, and the census journey-to-work data provides a unique opportunity for the analysis of at least one trip purpose.

2.2/ Prediction Versus Understanding

The transport systems models evolved during the late 1950s and early 1960s were responding to a transport policy environment that emphasized heavily the need to identify opportunities for long-range capital investments in new transport capacity. Rapidly growing urban areas and sharp increases in private car ownership and use stimulated this policy emphasis; therefore, the transport planning process was almost totally preoccupied with the prediction of future automobile volumes and proposed road networks.

This policy emphasis began to change during the late 1960s when many municipalities recognized that the so-called urban transport problem could not be solved simply by providing more road capacity. It was recognized that new investments in transport capacity would have to be accompanied by more enlightened urban development policies and the improved utilization of existing transport facilities. As a much wider range of public policy responses began to emerge, analytical tools that emphasized the prediction of future traffic volumes were of limited use. It seemed clear that an improved understanding of urban spatial structure should be developed to support this new policy environment. Many of the new policy initiatives were not only concerned with transport supply but also with the modification of transport demands. A major aim of the research projects on which this report is based was to develop an improved understanding of urban spatial structure and its relation to the journey to work.

3/ The Journey-To-Work Data

In the 1971 population and household census a *Place of Work* question was asked in every third private dwelling unit of each individual 15 years of age and older. The specific place of work questions asked in this sample of dwelling units are shown in Figure 2. For employed persons, the information desired was for the week prior to enumeration; other persons were asked to provide the information for the job of longest duration since January 1, 1970.

This place of work information has been coded to the municipal level for the entire country and to the census tract (CT) level for the 21 census metropolitan areas (CMAs) and 9 census agglomerations (CAs). A one-third sample of the data collected was coded and each observation weighted to approximate the total population. This means that the coded data represented approximately an 11% sample of the place of work of the resident labour force. It must be emphasized that throughout this report the analyses deal with home to work linkages and not trips as used in the conventional transport planning sense. The terms *trip generation* and *trip distribution* refer to trip linkages. Actual trips may be easily established by multiplying by an appropriate labour force based coefficient.

Figure 2/ 1971 Census of Canada Place of Work Questions

34. For whom did you work? **PLEASE PRINT:**

Name of firm,government,agency,etc.

Department,branch,division, or section

38. Where do you usually work?
Give this information for job described above. If no usual place of work, see Instruction Booklet.

Number **Street**

City,town,village or municipality **County** **Province**

O At home **O Outside Canada**

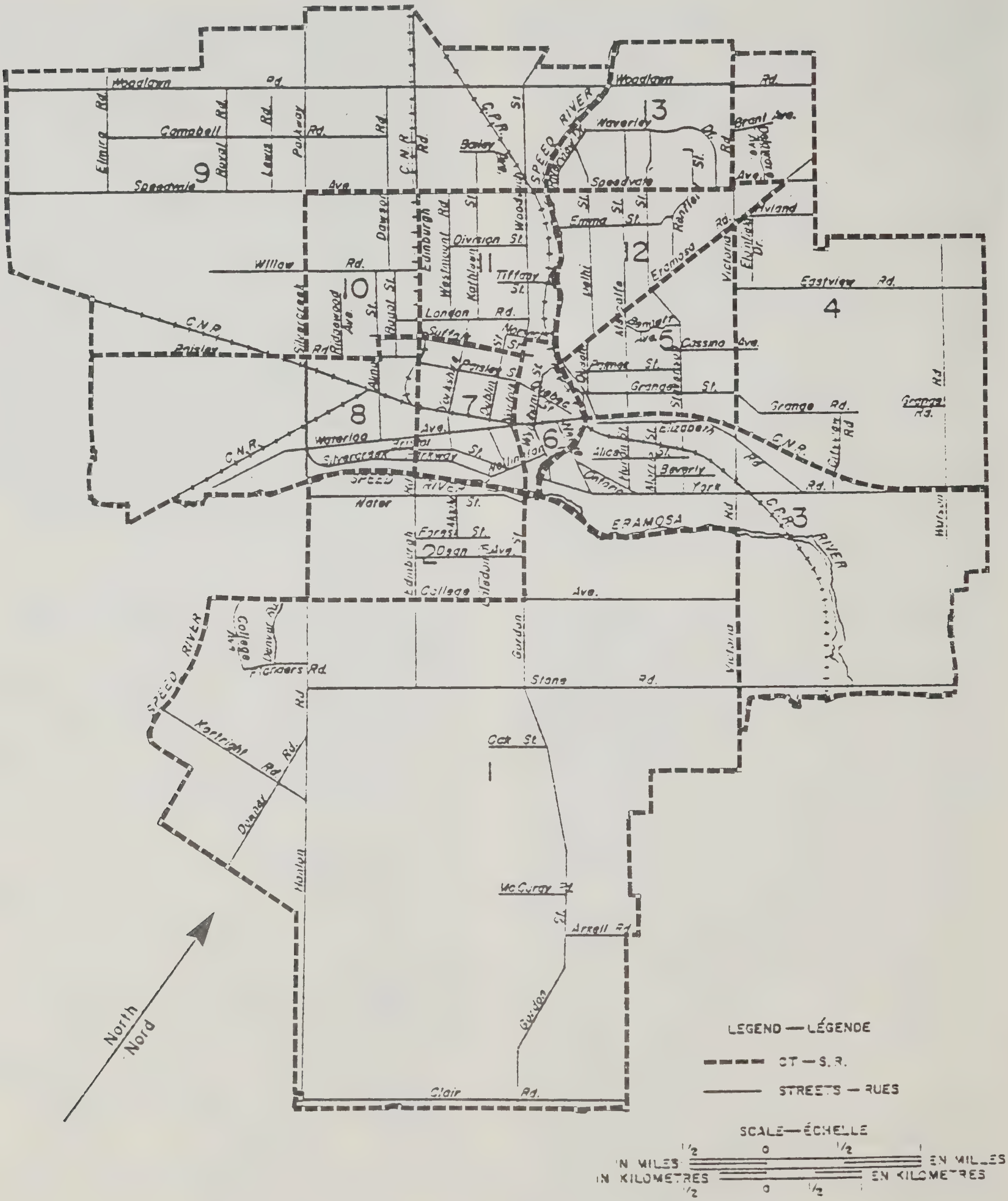


Figure 3/ 1971 Census Tract System for Guelph

3.1/ The Basic Spatial Unit

The basic spatial unit for coding the locations of the places of residence and work in the C(M)As is the census tract. Census tracts embrace the contiguous urbanized areas within a C(M)A. A distinction is made between official and unofficial census tracts. Unofficial census tracts cover areas just beginning to urbanize. A second spatial unit – the census division (CD) – is intended to cover those parts of an area not part of the contiguous urbanized area. In Ontario a census division corresponds to a county.

Statistics Canada coded the journey-to-work data to the census tract level if:

(i) a respondent's place of residence was located in a census tract within a C(M)A or a census district within 50 miles (80 km) of a C(M)A border; and
(ii) the respondent's place of work was located in a census tract.

Figure 3 shows the census tract system used for Guelph in the 1971 census. Fourteen census tracts were used along with 19 census districts. At the other end of the scale is the Toronto CMA, which in 1971 had 452 census tracts and 24 census districts.

3.2/ Format of Spatial Linkages Data

Figure 4 shows the data format for the home to work spatial linkages data. The rows identify the place of residence; the columns identify the place of work of employed residents. It should be noted from this table that complete place of residence by place of work flows exist only between census tracts. Flows from census tracts to census districts are not coded to specific census districts and inter-census district flows are not recorded.

The spatial linkages analyses described in this report are for the system of census tracts only. Census districts were not included because of the absence of complete data and the difficulty of rationally specifying the census tract centroids. Intra- and inter-census tract flows account for 75 to 85% of the journey-to-work flows in Ontario census areas as indicated in Table 1. Table 1 also indicates that trips from surrounding census districts to census tracts within the census areas represented from 3.0 to 18.4% of the total trips produced by the official census tracts. (As would be expected, census areas adjacent to larger metropolitan areas have lower proportions of internal trips as illustrated by Brantford and Oshawa.)

Figure 4/ Entries in Home to Work Linkage Matrix

PLACE OF RESIDENCE		PLACE OF WORK: CT WITHIN C(M)A 1 2 3 · · · · N	AT HOME	OUTSIDE C(M)A	OUTSIDE CANADA	CT NOT STATED	OTHER NOT STATED
CT WITHIN C(M)A	1						
	2						
	3						
	4						
	:						
	N						
CD WITHIN 50 MILE RADIUS	1						
	2						
	3						
	4						
	.						
	N						

3.3/ Road Network Coding

The census tract journey-to-work data have been supplemented with data on the road networks of each census area. Road networks have been coded at roughly the arterial/subarterial level using the 1971 census tract maps as a basis. Figure 5 shows the road network coded for Guelph and the way in which the census tract centroids have been linked to this network.

The length of each link has been measured in kilometres and minimum distance trees, and skim distances have been calculated for each census tract. Intra-census tract distances were obtained from Statistics Canada information on the weighted airline distances of home to work linkages between enumeration areas within each census tract. Consistent travel speed data were not available for the 16 areas. All the subsequent analyses described in this report are therefore based on road network distances rather than road network travel times. Provisions have been made in the coded networks to include link travel speeds if they become available at a later date.

3.4/ Special Tabulations

The Statistics Canada data tapes used for most of these analyses contain information on the home to work linkages of the employed labour force stratified by sex. Special tabulations were obtained from Statistics Canada for specific socio-economic groups to explore differences in behaviour between socio-economic groups. Three two-way classifications of the home to work matrices were useful for these special tabulations. Descriptions of these classifications follow.

Automobile Ownership – The "number of automobiles in a household" variable formed the basis for this classification. The two classes obtained are:

Group 1 – no automobile
- individuals from a household with 0 automobiles
- non-heads of household from a household with 1 automobile

Group 2 – automobile available
- heads of households from a 1 automobile household
- individual from households with 2+ automobiles

This grouping was used to approximate the linkage patterns of workers with and without access to a car for the journey to work.

Tenancy – The two groups used in this case are as follows:
Group 1 – owner occupied dwelling
Group 2 – rental dwelling.
The purpose of this grouping was to allow any differences in trip generation and trip distribution patterns to be detected.

Period of Residence – The length of time that the head of the household has resided at the particular address provided the basis for the following:
Group 1 – individuals residing at current dwelling for more than five years;
Group 2 – individuals residing at current dwelling for five years or less.
The purpose of this grouping was to try to isolate the spatial distribution characteristics of the newly relocated in an urban area.

3.5/ Spatial Units

Questions of confidentiality and statistical reliability arise when linkage matrices are requested that are stratified according to other census questions. For example, the Kitchener CMA contains 45 census tracts, but these census tracts have been aggregated into 25 zones for the special tabulations identified in the previous section. This aggregation of census tracts was organized so that the average flow between zones for each of the cross-classifications will be about 120 home to work linkages; this is similar to the magnitude when all census tracts are used. More specifically, the aggregation criteria were based on the following.

1/ The average inter-zonal flow should be about 120.
2/ The aggregated census tracts should share common boundaries.
3/ The aggregated zones should be reasonably uniform in labour force or employment size.
4/ The aggregated zones should be of a reasonably compact shape.
5/ There should be no natural boundaries cutting through the aggregated zones.

3.6/ Employment Data

Special tabulations of the spatial distribution of employment by industry sector were obtained from Statistics Canada. These employment data were grouped into 12 broad industry divisions, which were divided further into major groups. Three broad classes of employment were identified consisting of the following industry groups.

Basic Employment
Group 1 : Primary industries
Group 2 : Primary metal products, petroleum and coal products
Group 3 : Federal and provincial administration

Semi-Basic Employment
Group 4 : Miscellaneous employment including metal fabrication
Group 5 : Food, textiles, paper, wood, clothing
Group 6 : Construction
Group 7 : Transport, storage, warehousing, communications, utilities
Group 8 : Wholesale services
Group 12 : Business services

Service Employment
Group 9 : Retail services
Group 10 : Finance, insurance, real estate
Group 11 : Population services
Group 13 : Local government
Group 14 : Other

The totals for each industry group by census tract are obtained by summing those commuting only within the census area, so the employment magnitudes will be slightly less than actual employment levels.

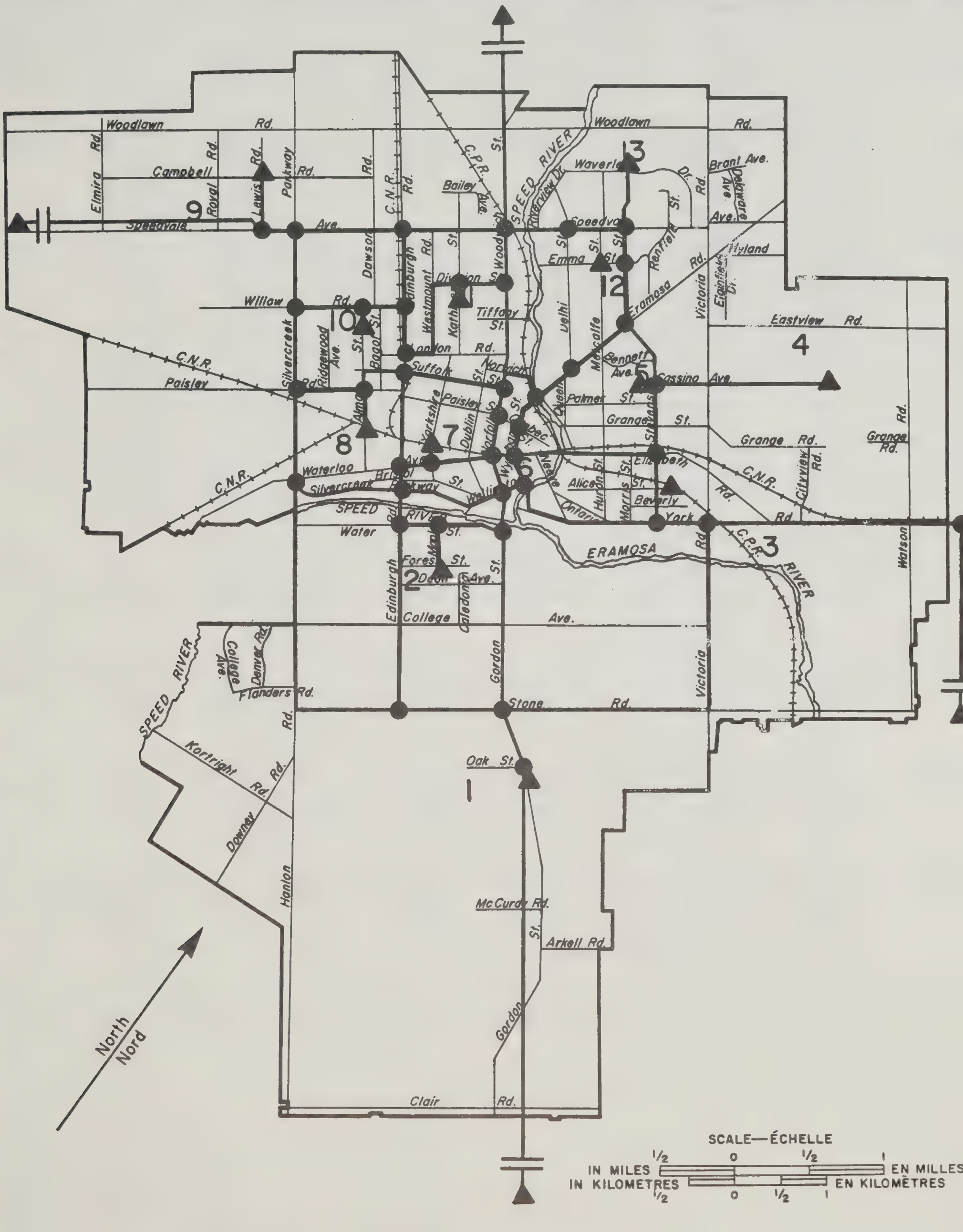


Figure 5/ Arterial Street System Coded for Guelph

Table 1/ Proportions of Various Work Trip Components for Ontario C(M)As

Census Area	Place of Resid.	CT in C(M)A	At Home	Outside C(M)A	Outside Canada	CT Not Stated	Other Not Stated	Total
Guelph	CT in C(M)A	19 990 (75.8%)	745 (2.8%)	2 515 (9.5%)	90 (0.3%)	1 020 (3.9%)	2 050 (2.8%)	26 365
	CD in 80 km	3 855(14.6%)** 3 555*						
Peterborough	CT in C(M)A	20 400 (80.5%)	950 (3.7%)	1 195 (4.7%)	150 (0.6%)	1 345 (5.3%)	1 630 (6.4%)	25 355
	CD in 80 km	2 520(5.5%) 2 425*						
Sarnia	CT in C(M)A	23 760 (78.7%)	1 670 (5.5%)	950 (3.1%)	250 (0.8%)	1 260 (4.2%)	2 245 (7.4%)	30 185
	CD in 80 km	2 520(8.3%) 2 425*						
Brantford	CT in C(M)A	24 195 (73.0%)	2 030 (6.1%)	2 410 (7.2%)	215 (0.6%)	340 (3.4%)	2 750 (8.3%)	33 125
	CD in 80 km	3 525(10.6%) 3 270*						
Sault Ste. Marie	CT in C(M)A	24 740 (80.7%)	1 570 (5.1%)	895 (2.9%)	120 (0.4%)	1 175 (3.8%)	2 070 (6.8%)	30 650
	CD in 80 km	1 385(4.5%) 1 370*						
Kingston	CT in (C(M)A	26 770 (76.5%)	1 340 (3.8%)	1 480 (4.2%)	300 (0.9%)	1 615 (4.6%)	3 100 (8.9%)	34 895
	CD in 80 km	5 440(15.6%) 5 315*						
Thunder Bay	CT in C(M)A	33 130 (76.3%)	1 270 (2.9%)	1 360 (7.1%)	245 (0.6%)	4 090 (9.4%)	3 180 (7.3%)	43 430
	CD in 80 km	825(1.9%) 800*						
Oshawa	CT in C(M)A	33 740 (72.7%)	1 165 (2.5%)	7 045 (15.2%)	150 (0.2%)	1 720 (3.7%)	2 780 (6.0%)	46 420
	CD in 80 km	8 550(18.4%) 8 470*						

Table 1/ Proportions of Various Work Trip Components for Ontario C(M)As (Cont'd)

Census Area	Place of Resid.	CT in C(M)A	At Home	Outside C(M)A	Outside Canada	CT Not Stated	Other Not Stated	Total
Sudbury	CT in C(M)A	45 580 (77.1%)	1 600 (2.7%)	2 890 (4.9%)	175 (0.2%)	3 310 (5.6%)	5 835 (9.9%)	59 145
	CD in 80 km	3 160(5.3%) 3 065*						
Kitchener	CT in C(M)A	81 095 (80.6%)	3 690 (3.7%)	4 710 (4.7%)	200 (0.2%)	2 990 (3.0%)	7 540 (7.5%)	100 595
	CD in 80 km	7 645(7.6%) 7 070*						
Windsor	CT in C(M)A	74 455 (76.5%)	4 330 (4.4%)	1 575 (1.6%)	4 765 (4.9%)	4 615 (4.7%)	7 415 (7.6%)	97 350
	CD in 80 km	4 745(4.9%) 4 695*						
London	CT in C(M)A	97 195 (79.6%)	6 650 (5.4%)	3 455 (2.8%)	520 (0.4%)	5 820 (4.8%)	9 180 (7.5%)	122 120
	CD in 80 km	6 540(5.4%) 6 125*						
St. Catharines	CT in C(M)A	91 815 (78.2%)	4 935 (4.2%)	4 070 (3.5%)	2 235 (1.9%)	5 225 (4.4%)	9 420 (8.0%)	117 445
	CD in 80 km	3 550(3.0%) 3 405*						
Hamilton	CT in C(M)A	152 265 (76.0%)	7 455 (3.7%)	14 960 (7.5%)	525 (0.3%)	7 940 (4.0%)	15 715 (7.9%)	199 810
	CD in 80 km	8 415(4.2%) 8 120*						
Ottawa	CT in C(M)A	201 015 (81.4%)	6 470 (2.6%)	3 255 (1.3%)	3 835 (1.6%)	12 020 (4.9%)	20 495 (8.3%)	246 905
	CD in 80 km	9 510(3.9%) 9 355*						
Toronto	CT in C(M)A	962 960 (81.9%)	40 975 (3.5%)	13 620 (1.2%)	2 985 (0.3%)	54 885 (4.7%)	100 340 (8.5%)	1 175 765
	CD in	34 685(3.0%)						

* Number of trips used from census divisions after dropping census divisions with trips to C(M)A less than 100.
** As a percentage of linkage produced by census tracts.

4/ The Ontario Census Areas

This section provides an overview of the characteristics of the Ontario census areas. At the 1971 census there were 16 census areas in Ontario, ranging in population size from Guelph to Toronto. The broad population and labour force characteristics of these census areas are examined in this section along with their broad spatial characteristics.

4.1/ Population Characteristics

Table 2 shows the population of the 1971 census areas and the number of census tracts and census districts within each. Six of the census areas had populations of less than 100 000; three had populations between 100 000 and 200 000; four had populations between 200 000 and 300 000; and three had populations of 500 000 or greater.

4.2/ Labour Force Composition

The 1971 labour force composition by industry division is available for each census area. Those industry divisions have been aggregated into eight industry groups listed in Table 3. Table 4 shows the proportions of the labour force in each census area that were within each of the eight industry groups. The last three rows of the table show the means, standard deviations and coefficients of variation of each of the proportions within each industry group for all census areas. The labour force proportions in Ontario and for all census areas are illustrated in Figure 6.

A review of Table 4 and Figure 6 shows that three industry groups account for about 70% of the total labour force: manufacturing, trade and services. The lowest coefficients of variation are for the construction, trade and services industry groups, suggesting that these are primarily populations serving industry groups and that the principal variations in the economic bases of the Ontario census areas may be found in the other industry groups.

Table 5 identifies those census areas with labour force proportions in each of the industry groups outside of the range of ±1 or 2 standard deviations from the mean value of an industry group. The entries in Table 5 may not be particularly surprising; however, they do provide some quantitative indication of the relative importance of the various industry groups to the economic bases of the Ontario census areas.

Table 6 identifies the dominant industry group in each Ontario census area on the basis of the information provided in Tables 4 and 5. Sault Ste. Marie,

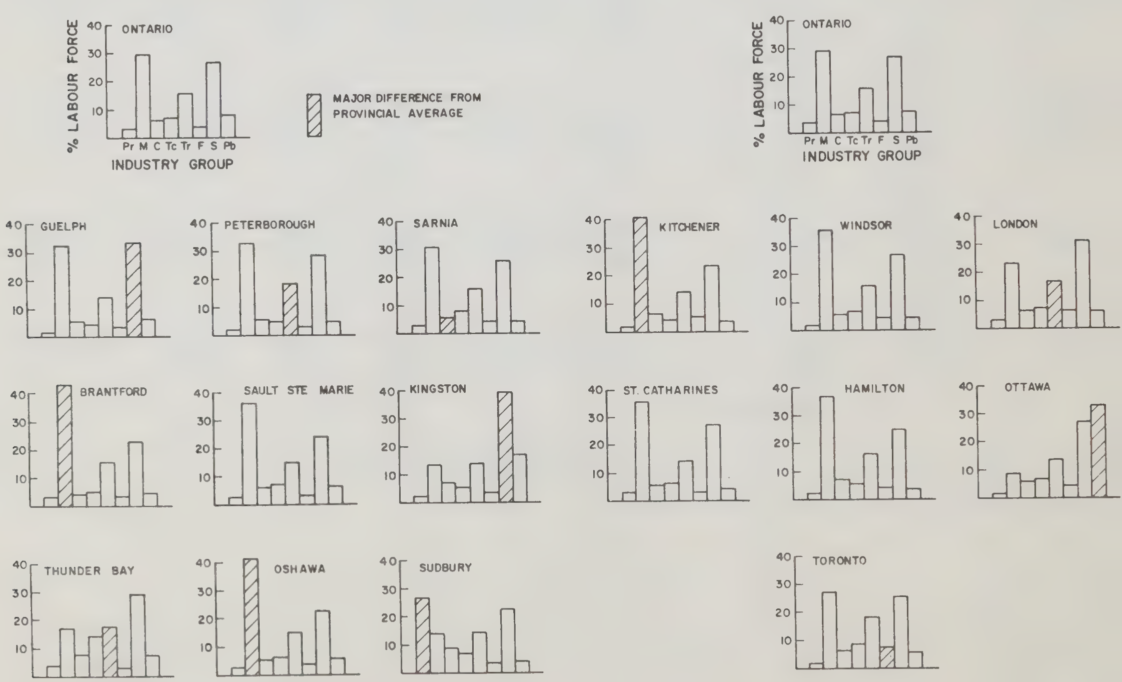


Figure 6/ Labour Force Proportions in Ontario Census Areas

Windsor, St. Catharines and Hamilton do not have dominant industry groups. Manufacturing industries provide the principal economic base for Brantford, Kitchener and Oshawa. The tertiary industry groups provide the principal economic base for Guelph, Peterborough, Kingston, London and Toronto. The transportation employment base of Thunder Bay, the mining base of Sudbury and the public service base of Ottawa are highlighted by Tables 4, 5 and 6. The extent to which the locational characteristics of these different industry groups influence the spatial properties of the census areas is examined later in this section.

4.3/ Geographic Characteristics

The amount of travel in an urban area is influenced by the geographic characteristics of an area. For example, average trip lengths would be expected to be longer in urban areas located on the Great Lakes because of the asymmetrical character of these areas. Another example is in the Hamilton area where the Hamilton Harbour and the Niagara Escarpment constrain urban development and travel.

It is difficult to express the topographic constraints that influence the spatial properties of urban areas in any simple way. Table 7 identifies the principal topographic characteristics of each Ontario census area and highlights those areas with major topographic constraints. Clearly, the major feature that may be used to group the areas is location, i.e., the urban area is inland or lakeside.

4.4/ Spatial Distributions of Households and Employment

Before examining the travel characteristics of the census areas in detail it is useful to highlight the major differences in the spatial distributions of households and jobs between the census areas. Again, no simple measure of spatial structure can be used to characterize each urban area. Figure 7 shows the spatial distribution of labour force and employment in the Kitchener census area. Figure 8 illustrates the distribution for London. The manufacturing employment base of the Kitchener census area is reflected in Figure 7 where employment is distributed widely across many census tracts. On the other hand, the concentration of employment in the London CBD reflects the strong financial employment base of the London CMA.

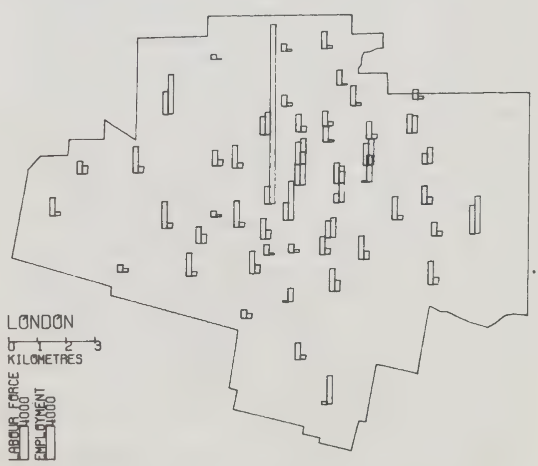
Table 8 summarizes the spatial properties of the Ontario census areas in 1971 using three broad groups of characteristics. The spatial characteristics of each census area are noted briefly in the following paragraphs.

Guelph - The employment was concentrated in three major areas: the CBD, the University of Guelph in the southern part of the area and the light industry area to the northwest. The dominant labour force proportions were in service and manufacturing, and these were reflected in the distribution of employment. Housing opportunities were well distributed allowing employees to live reasonably close to work. There are no major geographic constraints to distort the spatial distributions of households and jobs.

Peterborough - The dominant industry groups were trade, manufacturing and service; these groups were reflected in the concentration of employment in the CBD and to the southwest of the CBD. As in Guelph,

Figure 7/ Spatial Distribution of Labour Force and Employment in Kitchener CMA, 1971

Figure 8/ Spatial Distributions of Labour Force and Employment in London CMA, 1971



the housing opportunities were well distributed around the employment centres allowing employees to live close to work.

Sarnia - Unlike Guelph and Peterborough, manufacturing employment in Sarnia was concentrated in one census tract on the southwest edge of the area near the petrochemical complex. Other employment concentrations were located in the CBD and in the transport area to the north of the CBD. Residential areas were concentrated to the east and northeast away from the major job concentrations, which created longer trip lengths to work. Longer trip lengths were also created by the semi-circular pattern of urbanization along the St. Clair River and Lake Huron.

Brantford - Brantford had the highest proportion of manufacturing employment of any Ontario census area. Employment was distributed throughout the area, with significant concentrations in six census tracts. Housing opportunities were well distributed throughout the area, providing a good spatial balance between households and jobs. There were no significant geographic constraints to the settlement patterns except the Grand River flood plain.

Sault Ste. Marie - There were no dominant housing groups relative to the provincial averages but there was an above average concentration in manufacturing. Most of this employment was concentrated in the steel mills in the western part of the area on the St. Mary's River. The residential areas were concentrated largely to the east and northeast of the employment concentrations, resulting in some longer trip lengths. The location along the St. Mary's River is a significant geographic constraint to development.

Kingston - Fifty-four percent of the labour force was concentrated in service and public administration employment, reflecting the presence of Queen's University, the hospitals, the prisons and military establishments. Employment concentrations occurred in five census tracts (four in Kingston and one in Kingston Township). Many newer residential areas were located in Kingston Township. This created long trips to the jobs in Kingston. Development and commuting patterns were heavily influenced by Kingston's lakeside location and the voids in development created by prison lands and the marsh lands of the Cataraqui River.

Thunder Bay - The dominant industry group was transportation and communications, which created nodes of employment. However, employment was well distributed throughout the area. Households were also well distributed, with some tendency to concentrations in the southwest. The location of Thunder Bay on Lake Superior and certain topographic constraints influence urban development patterns.

Oshawa - The dominant employment was manufacturing, with much of this concentrated in the automobile assembly plant. The CBD also had a significant concentration of employment. In spite of Oshawa's location on Lake Ontario, these two employment zones are located away from the shoreline. Although employment was concentrated, the residential areas were well distributed throughout the area and around these employment concentrations.

Sudbury - The location of mining employment was the dominant feature of the spatial structure. There was also a significant concentration of employment in the CBD. Housing opportunities were well distributed throughout the area. The Canadian Shield topography influences settlement patterns.

Kitchener - Employment was concentrated in the manufacturing industry groups with no dominant industrial establishments. Employment opportunities were widely distributed throughout the Kitchener-Waterloo and Cambridge areas. Housing opportunities were also well distributed, allowing employees to live close to work. The Grand River and its flood plain represent the major geographic constraints to development in the area.

Windsor - None of the employment groups was particularly dominant, although manufacturing was above the provincial average due to the presence of automotive industries. Employment was concentrated in about five census tracts, with housing opportunities well distributed throughout the area around each of the employment concentrations. The riverside and lakeside location of Windsor influences the patterns of urbanization.

London - The service employment industry group dominated the employment, much of which was due to the presence of the University of Western Ontario. Employment was concentrated in the CBD and to the southwest, with the CBD having a relatively heavy concentration. The only major geographic constraint is the Thames River.

St. Catharines - This census area consists of the relatively self-contained communities of St. Catharines, Niagara Falls and Welland. No particular industry group dominated the employment base of the area, although employment in manufacturing was above the provincial average. Employment and housing opportunities were well distributed, allowing short trip lengths. The Lake Ontario shoreline, the Niagara River and the Welland Canal do constrain settlement patterns.

Hamilton - Although no industry group dominated the employment base, the Hamilton census area did have above average concentrations of employment in the manufacturing and transportation and communications groups. Because of the large-scale industries in Hamilton, employment was concentrated along the Harbour. The CBD also contained a significant concentration of employment, while housing opportunities were concentrated to the south both below and above the Niagara Escarpment and in Burlington. Patterns of urbanization in Hamilton are heavily influenced by Lake Ontario, the Harbour and the Escarpment.

Ottawa - Much of the public sector employment was concentrated in three census tracts in the centre of Ottawa along with service employment. There are other concentrations of employment throughout the Ottawa and Hull areas. Housing opportunities were well distributed throughout the area. The Ottawa and Rideau Rivers constrain development.

Toronto - Employment was concentrated in the trade and service industry groups. Much of the employment in these sectors was concentrated in the central area zones. Manufacturing employment was concentrated in the older industrial areas and along the freeway network. Although housing was well distributed throughout the census area, there was a concentration of jobs in Metropolitan Toronto.

4.5/ Population Growth 1971-1976

The principal concern of this report is the 1971 census data; however, it is useful to examine the population changes in the Ontario census areas between 1971 and 1976. Table 9 shows the 1971 and 1976 populations recorded by Statistics Canada for the 16 Ontario census areas with their growth rates for 1966-1971 and 1971-1976. The populations in this table are for the same area as the 1976 population figures and in general are different from those in Table 2. It should be noted that the 1971-1976 five-year growth rate varied from 13% in the Kitchener CMA to -2% in the Windsor CMA . The population growth rates during 1971-1976 were all lower than during 1966-1971 with some dramatic changes, such as in Sault Ste. Marie, Sudbury, Windsor and Toronto.

Changes in the spatial distributions of population by census tract are also of interest in order to determine whether the 1971 spatial linkages data might still be applicable. The spatial changes have been explored in five census areas: Brantford, Sault Ste. Marie, Oshawa, Kitchener and Windsor. Figures 9 to 13 show the 1976 census tract populations plotted against the 1971 census tract populations for the five cities and the spatial distributions of the positive and negative changes in census tract population levels.

The five-year population growth rate in Brantford was only 2.2% during 1971-1976. CT 14 in the north end of the area was the only census tract to experience significant population growth; its population increased from 7 300 to about 12 900. The largest decline was in CT 1 where the population decreased from about 9 200 to about 9 000.

The five-year population growth rate in Sault Ste. Marie was negligible during 1971-1976 but there were some shifts in the spatial distributions of households. Census tracts around the steel mill area lost population with most of the growth occurring in the east and northwest. The largest decrease was 1 400 in CT 7; the largest increase was 2 200 in CT 1 located on the eastern boundary of the area.

The five-year population growth was 11% in Oshawa during 1971-1976 and most of this growth occurred in census tracts on the edges of the area. The largest increases were in CT 9, 13, 14 and 15 to the north and CT 2 on Lake Ontario. The largest decrease was 900 in CT 7, which is adjacent to the CBD.

The Kitchener census area had the highest population growth rate during 1971-1976 (13%). Most of the growth occurred in CT 2, 8 and 9 in the south and west of Kitchener and CT 101 in the west of Waterloo. Figure 12 demonstrates that most of the population decreases occurred in the inner census tracts. The largest decrease occurred in CT 21, which decreased by 900.

The Windsor CMA experienced a net loss in population during 1971-1976. Figure 13 shows that all census tracts lost population except several census tracts located on the fringes of the area. Population increases of 4 200 and 7 200 occurred in CT 18 and 19, respectively, while CT 1 experienced the largest decline in population (about 1 000).

The evidence presented for the five census areas demonstrates that the only significant changes in the spatial distributions of households have occurred on the fringes of the urban areas where new residential subdivisions have developed. While there have been some decreases in the populations of the central areas, most of this decrease seems to be due to changing household structure rather than to the migration of households.

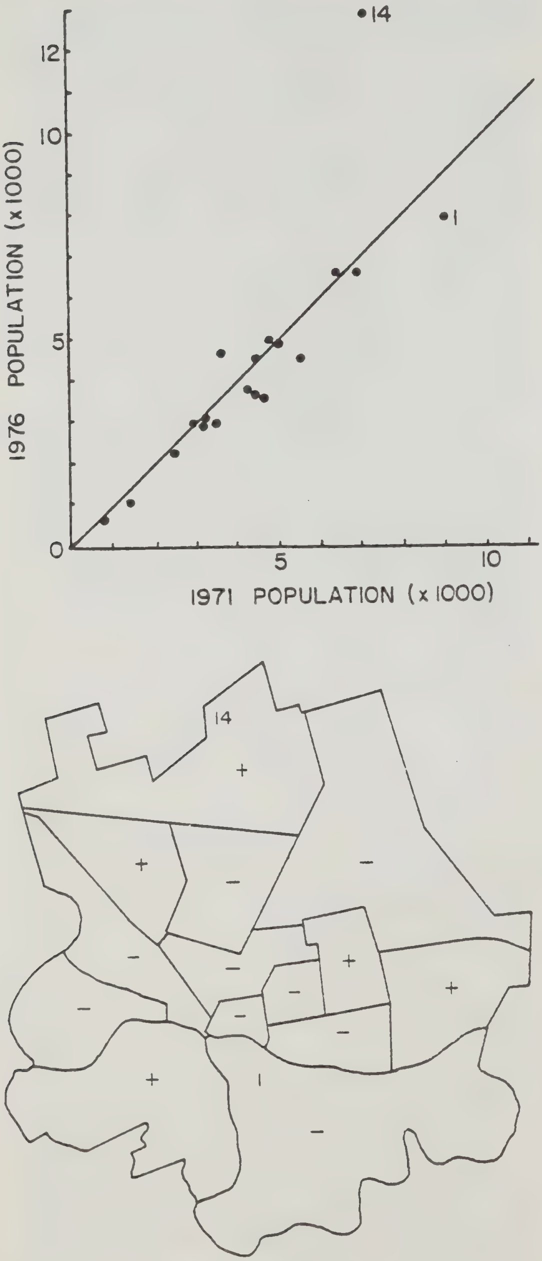


Figure 9/ Census Tract Population Changes in Brantford 1971-1976

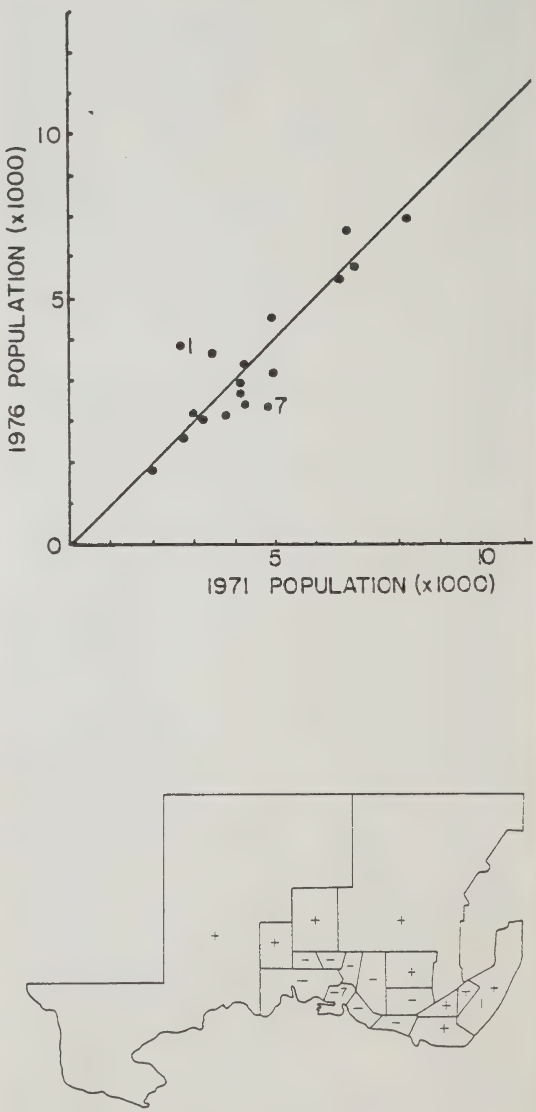


Figure 10/ Census Tract Population Changes in Sault Ste. Marie 1971-1976

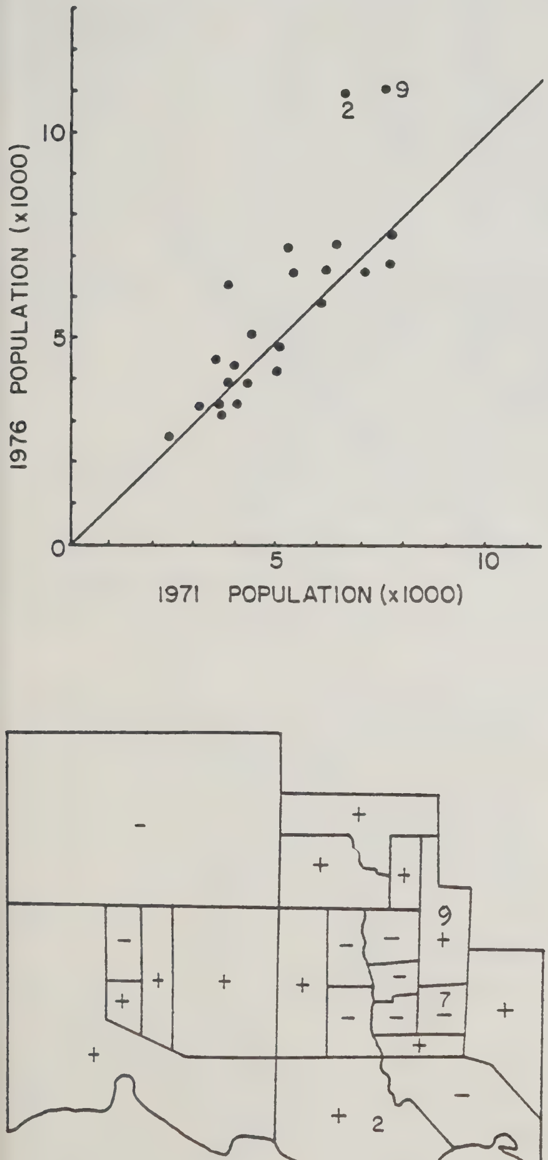


Figure 11/ Census Tract Population Changes in Oshawa 1971-1976

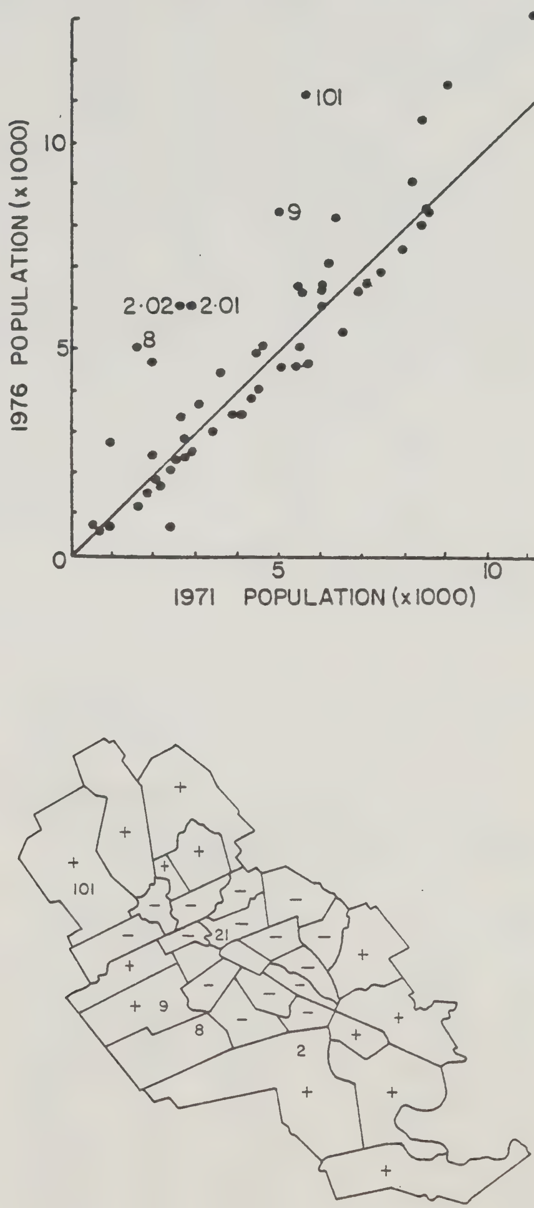
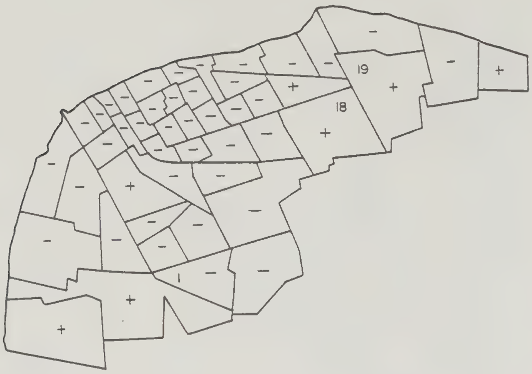


Figure 12/ Census Tract Population Changes in Kitchener 1971-1976



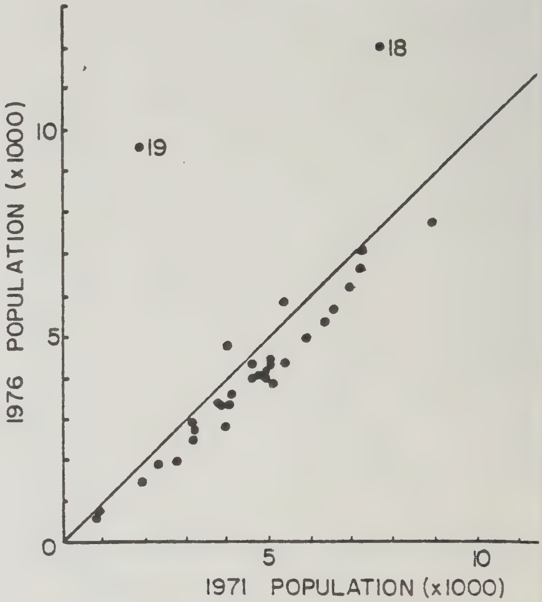


Figure 13/ Census Tract Population Changes in Windsor 1971–1976

Table 2/ 1971 Census of Canada: Ontario Census Metropolitan Areas and Census Agglomerations

C(M)A Number	Census Area	1971 Population	Number of CT's	Number of CD's
1	Guelph	62 660	14	19
2	Peterborough	63 530	16	10
3	Sarnia	78 445	18	6
4	Brantford	80 285	16	16
5	Sault Ste. Marie	81 270	19	2
6	Kingston	85 875	18	8
7	Thunder Bay	112 095	25	2
8	Oshawa	120 320	22	10
9	Sudbury	155 425	28	5
10	Kitchener	226 850	45	21
11	Windsor	258 645	56	5
12	London	286 010	59	15
13	St. Catharines	303 430	53	11
14	Hamilton	498 525	109	19
15	Ottawa	602 510	120	19
16	Toronto	2 628 045	452	24

Table 3/ Census Industry Division Groups

Group	Code	Census Industry Divisions	Type
1	Pr	1,2,3,4	Primary, Agriculture, Forestry, Fishing & Trapping, Mines, Quarries & Oil Wells
2	M	5	Manufacturing
3	C	6	Construction
4	Tc	7	Transportation, Communications & Other Utilities
5	Tr	8	Trade
6	F	9	Finance, Insurance, Real Estate
7	S	10	Community, Business and Personal Service
8	Pb	11	Public Administration & Defence

Table 4/ Proportions of Labour Force by Industry Group

Census Area	Industry Group Pr	M	C	Tc	Tr	F	S	Pb
Guelph	1.7	31.7	5.8	4.0	14.1	3.3	32.8	6.6
Peterborough	1.1	32.5	5.6	5.1	18.5	3.8	28.4	5.0
Sarnia	2.6	30.9	5.7	8.0	15.7	4.4	25.4	4.3
Brantford	3.2	42.7	4.4	4.6	15.6	2.9	22.2	4.4
Sault Ste. Marie	1.8	36.7	6.2	6.9	15.0	2.6	24.3	6.5
Kingston	1.4	13.1	6.7	5.1	13.9	3.7	39.8	16.3
Thunder Bay	3.8	17.1	7.6	14.8	17.3	3.2	29.2	7.0
Oshawa	1.5	41.5	5.2	5.7	15.0	3.6	22.4	5.1
Sudbury	26.5	14.3	8.8	6.2	14.1	3.0	22.3	4.8
Kitchener	1.2	41.5	6.5	4.1	14.7	5.1	23.2	3.7
Windsor	1.7	35.5	5.5	6.5	15.9	4.3	26.4	4.4
London	3.1	23.7	6.3	6.6	16.4	6.6	30.7	6.6
St. Catharines	2.8	35.5	6.0	6.7	14.5	3.2	26.8	4.5
Hamilton	2.0	36.9	6.7	5.2	16.2	4.1	24.9	4.0
Ottawa	1.0	8.6	6.3	6.7	13.2	4.7	26.9	32.6
Toronto	1.0	27.3	6.5	8.1	18.0	7.3	26.0	5.8
Mean	3.53	29.36	6.43	6.51	15.52	4.11	26.99	7.60
Standard Deviation	6.21	10.92	1.16	2.53	1.49	1.31	4.57	7.29
Coeff. of Variation	1.76	0.37	0.18	0.39	0.10	0.32	0.17	0.96

Table 5/ Census Areas with Atypical Proportions of Labour Force by Industry Group

Industry Group	Census Areas with Labour Force Proportions Outside Mean + S.D.	Mean – S.D.
Primary	Sudbury*	
Manufacturing	Brantford Oshawa Kitchener	Kingston Thunder Bay Sudbury Ottawa
Construction	Sarnia* Sudbury*	Brantford
Transportation	Thunder Bay*	
Trade	Peterborough Thunder Bay Toronto	Ottawa
Finance	London Toronto*	
Service	Kingston* Guelph	
Public Adminstration	Kingston Ottawa*	

*Greater than mean + two standard deviations

Table 6/ Dominant Industry Groups in Each Census Area

Census Area	Dominant Industry Group P	M	C	Tc	Tr	F	S	Pb
Guelph							*	
Peterborough					*			
Sarnia			*					
Brantford		*						
Sault Ste. Marie								
Kingston							*	*
Thunder Bay				*				
Oshawa		*						
Sudbury	*							
Kitchener		*						
Windsor								
London						*		
St. Catharines								
Hamilton								
Ottawa								*
Toronto					*	*		

Table 7/ Principal Topographic Characteristics of the Census Area

Census Area	Topographic Characteristics	Major Constraint
Guelph	No strong topographic constraint except trisection by Speed and Eramosa Rivers	
Peterborough	Bisection by Otonabee River and Trent Canal	
Sarnia	Location on St. Clair River and Lake Huron	X
Brantford	Grand River passes through southwest sector	
Sault Ste. Marie	Located on St. Mary's River	X
Kingston	Located on Lake Ontario and the Great Cataraqui River and cut by the Little Cataraqui River	X
Thunder Bay	Located on Lake Superior and bounded on south by Taministikwia River and mountain	X
Oshawa	Located on Lake Ontario	
Sudbury	Canadian Shield topography	
Kitchener	Bisected by Grand River	
Windsor	Location on Detroit River and Lake St. Clair	X
London	Trisection by Thames River	
St. Catharines	St. Catharines located on Lake Ontario, Niagara Falls located on Niagara River and Welland Canal bisects area	X
Hamilton	Lake Ontario location and presence of Hamilton Harbour and Niagara Escarpment to south	X
Ottawa	Area bisected by Ottawa River and Ottawa cut by Rideau River and Canal	X
Toronto	Lake Ontario Location	X

Table 8/ Spatial Structure of Census Area

	Spatial Structure			Employment		Households	
Census Area	Concentric	Semi Circular	Several Component Cities	Concentrated	Dispersed	In Balance with Jobs	Away From Job Centres
Guelph	X				X	X	
Peterborough	X				X	X	
Sarnia		X		X			X
Brantford	X				X	X	
Sault Ste. Marie		X		X			X
Kingston		X			X		X
Thunder Bay		X	X		X	X	
Oshawa	X			X		X	
Sudbury	X			X			X
Kitchener	X		X		X	X	
Windsor		X		X		X	
London	X				X	X	
St. Catharines	X		X		X	X	
Hamilton		X	X	X			X
Ottawa	X		X	X		X	
Toronto		X	X	X			X

5/ Aggregate Travel Demands

Table 9/ Ontario Census Area Population 1971-1976

Census Area	1966-1971 Growth[1]	Population 1971[2]	1976	1971-1976 Growth[3]
Guelph	16.7	62 659	70 148	12.0
Peterborough	3.7	63 531	64 618	1.7
Sarnia	5.6	78 444	80 387	2.5
Brantford	6.8	80 292	82 080	2.2
Sault Ste. Marie	7.6	81 270	81 375	0.1
Kingston	4.5	85 877	89 532	4.1
Thunder Bay	3.8	114 708	117 988	2.9
Oshawa	13.0	120 318	133 959	11.3
Sudbury	13.7	157 721	155 013	-1.7
Kitchener	18.0	238 574	269 828	13.1
Windsor	8.5	248 718	243 319	-2.2
London	12.7	252 981	264 639	4.6
St. Catharines	6.3	285 802	298 129	4.3
Hamilton	9.0	503 122	525 222	4.4
Ottawa	14.0	619 861	672 166	8.4
Toronto	14.8	2 602 098	2 753 082	5.8

[1] Percentage of 1966 population

[2] Population of same area as 1976 census

[3] Percentage of 1971 census

The initial analyses of the census journey-to-work data were concerned with the total amount of home to work travel in each census area and the relationship of this travel to the spatial organization of the area. This aggregate analysis of commuting provides a background for the more detailed analyses of spatial interaction described in subsequent sections. Clearly, the total amount of travel between households and jobs in an area depends on the spatial arrangements of these activities, the average home to work trip length and the labour participation rate.

The aggregate amount of travel between households and jobs has been calculated for each census area by combining the census home to work flows with the minimum inter-census tract distances obtained from the coded road networks. Table 10 shows a number of indicators of home to work travel for each Ontario census area. The total movement between homes and jobs located in urban area census tracts varies from about 73 000 km in Peterborough to about 10 300 000 km in the Toronto CMA. Per capita movement varies from 1.16 km in Peterborough to 3.93 km in Toronto. The two right-hand columns of Table 10 show that this variation in the amount of travel is due to variations in the mean trip distance as well as to variations in the labour participation rate.

Table 10/ Spatially Aggregated Indicators of Travel Demand

	Census Area	Population	Total Movement (km)	Km Per Capita	Mean Trip Length (km)	Labour Participation Rate
1	Guelph	62 660	77 933	1.24	3.9	0.403
2	Peterborough	63 530	73 436	1.16	3.6	0.407
3	Sarnia	78 945	154 472	1.96	6.5	0.396
4	Brantford	80 285	106 458	1.33	4.4	0.406
5	Sault Ste. Marie	81 270	123 695	1.52	5.0	0.404
6	Kingston	85 875	147 251	1.71	5.5	0.422
7	Thunder Bay	112 095	205 331	1.83	6.2	0.409
8	Oshawa	120 320	151 884	1.26	4.5	0.376
9	Sudbury	155 425	401 068	2.58	8.8	0.450
10	Kitchener	226 850	421 683	1.86	5.2	0.428
11	Windsor	258 645	595 784	2.30	8.0	0.392
12	London	286 010	680 267	2.38	7.0	0.416
13	St. Catharines	303 430	670 235	2.21	7.3	0.379
14	Hamilton	498 525	1 278 598	2.56	8.4	0.410
15	Ottawa	602 510	1 817 155	3.02	8.3	0.427
16	Toronto	2 628 045	10 332 392	3.93	10.8	0.450

5.1/ Per Capita Movement and Census Area Population

Figure 14 shows the relationship between the per capita amount of travel and the population of each census area. This diagram shows that the per capita home to work distance increases with increasing census area population but at a decreasing rate. While the observations tend to cluster around this trend line, there are some significant deviations from it. For example, both Sarnia and Sudbury had per capita travel amounts significantly above the average, while Oshawa and the Kitchener census areas had per capita travel amounts less than the average.

The variations in the amount of travel illustrated in Figure 14 are due to variations in the mean trip length and to variations in the labour participation rate. The following regression equation was fitted to the date in order to determine the relative importance of the two effects:

$$\text{km/capita} = -2.229 + 0.324 \cdot \text{mean trip length} + 5.327 \cdot \text{labour participation rate} \qquad (2)$$

$R^2 = 0.95 \qquad \beta_1 = 0.89 \qquad \beta_2 = 0.15$

Equation 2 demonstrates that while the labour force participation rate explains some of the observed variation in home to work travel, the dominant variable is the mean trip length.

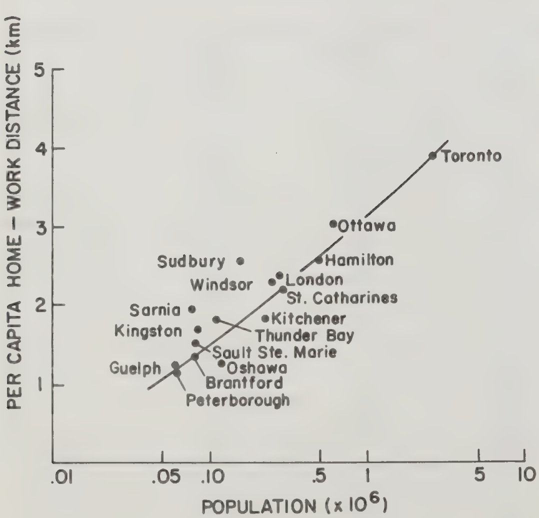


Figure 14/ Per Capita Home-to-Work Distance vs Census Area Population

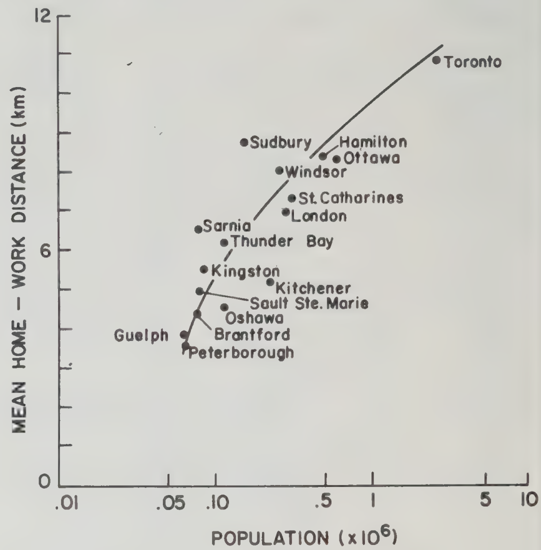


Figure 15/ Mean Home-to-Work Distance vs Census Area Population

Figure 15 shows the relationship between the mean home to work distance and population for the 16 Ontario census areas. This diagram contains essentially the same information as Figure 14; however, it highlights the variations in mean trip distance. The mean trip distance is the variable that is central to the calibration of the spatial interaction models described later in this report. The wide range in mean trip length for the census areas with populations of 150 000 and less should be noted.

5.2/Aggregate Travel and Spatial Structure

In discussing the impact of spatial structure on aggregate home-work travel, it is useful to group the Ontario census areas into five broad groups to examine differences in aggregate travel. These groups are as follows:

A. Guelph, Peterborough, Sarnia, Brantford, Sault Ste. Marie and Kingston
B. Thunder Bay, Oshawa and Sudbury
C. Kitchener, Windsor, London and St. Catharines
D. Hamilton and Ottawa
E. Toronto

Group A - Figure 15 shows that the mean trip lengths of the Group A census areas varied from 3.6 km in Peterborough to 6.5 km in Sarnia. Guelph, Peterborough and Brantford have the lowest mean trip lengths. It was pointed out in Section 3 that there are few geographic constraints to development in these areas and that households and jobs are well dis-

tributed. In contrast, the other three Group A communities have shoreline locations and significant job concentrations in one or two census tracts along these shorelines. Longer trip lengths are created by these geographic constraints as well as by the spatial imbalance between households and jobs. The differences in trip length characteristics of the two subgroups are evident in Figure 16, which illustrates the mean trip length contours from the households in Peterborough and Sarnia. The map for Peterborough shows that there is a large area of the central city in which the mean trip length is less than 3.0 km. In contrast, the long trip lengths from the northern section of Sarnia to the employment concentration in the Petrochemical complex to the southwest are well illustrated.

Group B - Figure 15 shows that the mean trip lengths of the Group B census areas were 4.5 km in Oshawa, 6.2 km in Thunder Bay and 8.0 km in Sudbury. It was pointed out in Section 3 that while Oshawa is located on Lake Ontario and has a concentrated distribution of employment, households are well distributed throughout the area. A second factor contributing to the lower mean trip length in Oshawa is the amount of commuting to jobs outside of the census area. This is not included in the calculation of the mean trip lengths presented in Table 10. Table 1 showed that in 1971 approximately 15% of Oshawa residents travelled to jobs outside of the census area. While Thunder Bay has some significant geographic constraints to development because of its location on Lake Superior, jobs and households are well distributed throughout the area, and in 1971 there was a degree of localization of commuting patterns within each of the former communities of Fort William and Port Arthur. The long trip lengths due to the location of mining employment in Sudbury were already mentioned in Section 2.

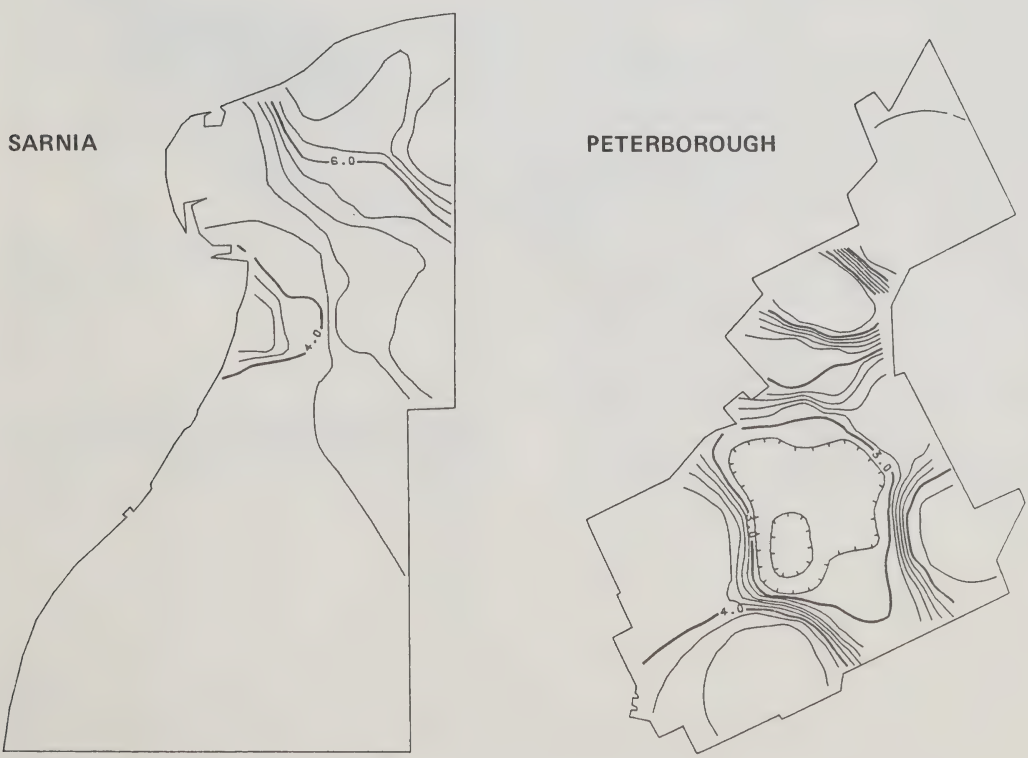


Figure 16/ Mean Trip Length Contours From Households in Sarnia and Peterborough

Group C – Average trip lengths in the four Group C communities ranged from 5.2 km in the Kitchener CMA to 8.0 km in Windsor. The lower mean trip length in the Kitchener CMA was due to the broad distribution of households and jobs in the area and to the localization of commuting within the Kitchener-Waterloo and Cambridge areas. There was also a similar degree of self-containment within the St. Catharines census area in the cities of St. Catharines, Niagara Falls and Welland. While there was strong concentration of jobs in the London CBD in 1971, long trip lengths did not occur because housing opportunities are well distributed around the CBD and other employment centres, and there were no significant geographic constraints to travel.

The mean trip length in Windsor was longer than in the other three census areas in this group because of the relative concentration of households in the northeastern part of the area away from the major job concentrations. This influence is evident in Figure 17, which shows the spatial distributions of labour force and employment and the commuting patterns in the Windsor CMA.

Group D – The mean trip length in Hamilton and Ottawa were similar in 1971. Jobs in both Hamilton and Ottawa were concentrated in the central areas. Longer trip lengths to these job concentrations resulted from the geographic constraints to development in both areas. The mean trip length in Ottawa was marginally lower than that in Hamilton because of the tendency of self-containment within the Ottawa and Hull areas.

Group E – Travel within the Toronto CMA cannot be realistically compared with travel in the other Ontario census areas because of the large population and geographic scale of the census area. Table 10 demonstrates that while the population of the Toronto CMA was some five times larger than the population of Ottawa and Hamilton CMAs, the mean trip length was only about 25% higher. The wide distribution of jobs throughout the census area allows many local commutersheds to develop. Figure 18 illustrates some of the major commuter flows within the Toronto CMA.

Figure 17/ Labour Force and Employment Distribution and Commuting Pattern in Windsor, 1971

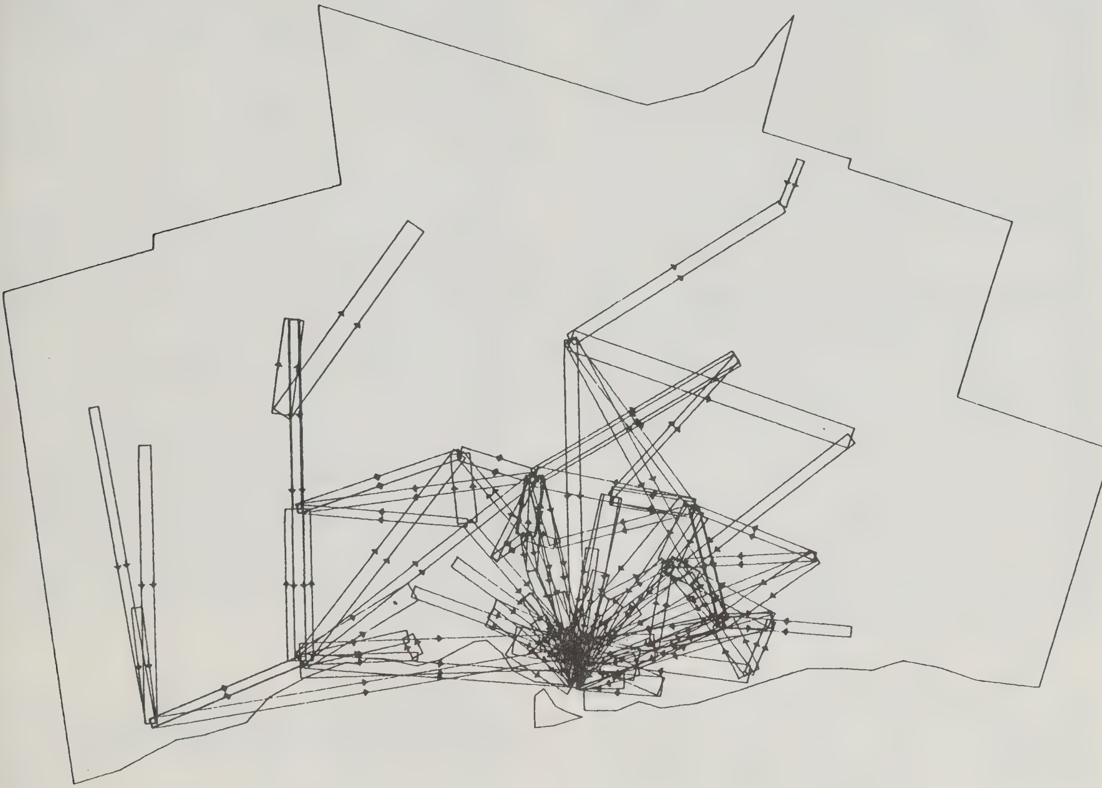


Figure 18/ Major Commuter Flows in Toronto, 1971

6/ Census Tract Labour Force Prediction Equations

The initial step in estimating travel demand is trip generation, in which the land use predicted for each traffic analysis zone is converted into trip ends. This section analyses the variations in the labour force composition of census tracts and relates this variation to several measures of residential activity. It was pointed out previously that work trip production rates cannot be derived directly from the census data, but only from home to work linkage rates. These rates may be treated as trip production rates if one assumes that each employed member of the labour force made one journey to work on the day of the survey.

6.1/ Aggregate Prediction Equations for Population and Households

Table 11 shows the set of simple regression equations that were developed between census tract labour force and census tract population, and between census tract labour force and census tract households. This table illustrates that both the census tract population and the census tract households explain much of the observed variation in census tract labour force. The coefficients of determination for the regression equation using population as the independent variable varies from 0.87 in Ottawa to 0.99 in Peterborough, and for the household regression equation, varies from 0.88 in London and Hamilton to 0.98 in Brantford and Sault Ste. Marie. The regression equations are of high statistical quality; however, some of the constant terms are rather large. In addition, inspection of the residuals of these equations suggested that there is a spatial bias to the residuals in several areas that is associated with variation in dwelling unit composition.

6.2/ Census Tract Labour Force versus Dwelling Unit Composition

Table 12 shows the set of multiple regression equations developed using the three classes of dwelling units identified in the household census. Table 12 shows that the coefficient of determination is relatively high for all the census areas, varying between 0.92 and 0.99. An examination of the statistical significance of the partial regression coefficients showed that in many cases the coefficient b_2 of the number of single attached dwelling units was marginally significant in Guelph, Brantford and Sudbury and not significant in the St. Catharines CMA.

Inspection of the b_1 magnitudes in Table 12 shows that they vary from 1.444 in Thunder Bay to 1.723 in the Kitchener CMA, with the majority of the magnitudes falling between 1.5 and 1.6 labour force per single detached dwelling unit. If the non-significant magnitudes of b_2 are omitted, then b_2 varies from 0.946 in Windsor to 2.115 in Sault Ste. Marie. The magnitudes of b_3 vary from 0.737 in Peterborough to 1.466 in Sudbury.

Table 11/ Census Tract Labour Force Prediction Equations

Census Area	Population[1]			Households[2]		
	a	b	R^2	a	b	R^2
Guelph	86	0.403	0.94	-1	1.443	0.94
Peterborough	-24	0.407	0.99	72	1.290	0.91
Sarnia	-33	0.396	0.97	1	1.321	0.90
Brantford	21	0.406	0.98	13	1.354	0.98
Sault Ste. Marie	-117	0.404	0.95	3	1.455	0.98
Kingston	-18	0.422	0.92	97	1.265	0.95
Thunder Bay	-105	0.409	0.97	12	1.332	0.94
Oshawa	63	0.375	0.94	-10	1.375	0.94
Sudbury	-220	0.450	0.90	-20	1.526	0.95
Kitchener	83	0.428	0.98	-106	1.570	0.97
Windsor	-50	0.392	0.96	-7	1.313	0.89
London	61	0.416	0.95	260	1.197	0.88
St. Catharines	39	0.378	0.99	54	1.282	0.97
Hamilton	-25	0.410	0.92	98	1.268	0.88
Ottawa	-37	0.427	0.87	221	1.295	0.89
Toronto	-17	0.450	0.91	138	1.439	0.89

[1] C.T. Labour Force = a + b × C.T. Population

[2] C.T. Labour Force = a + b × C.T. Households

Table 12/ Census Tract Labour Force Prediction Equation Using Dwelling Unit Composition

Census Area	a	b_1	b_2	b_3	R^2
Guelph	30	1.570	0.902	1.262	0.96
Peterborough	1	1.525	1.639	0.737	0.98
Sarnia	19	1.490	1.994	0.685	0.98
Brantford	42	1.500	1.214	0.965	0.98
Sault Ste. Marie	-34	1.563	2.115	1.152	0.98
Kingston	20	1.534	1.266	1.136	0.98
Thunder Bay	19	1.444	1.448	0.870	0.96
Oshawa	-30	1.508	1.505	1.037	0.97
Sudbury	-60	1.791	0.431	1.466	0.97
Kitchener	-88	1.723	1.911	1.206	0.98
Windsor	-69	1.575	0.946	0.889	0.97
London	145	1.474	1.895	0.923	0.92
St. Catharines	15	1.511	0.625	0.849	0.99
Hamilton	-32	1.552	1.414	1.054	0.92
Ottawa	41	1.644	1.650	1.112	0.95
Toronto	16	1.574	1.732	1.324	0.96

C.T. Labour Force = a + b_1 × C.T. Single Detached Dwelling Units
+ b_2 × C.T. Single Attached Dwelling Units
+ b_3 × C.T. Apartments

One of the difficulties in establishing reliable estimates of b_2 and b_3 is that the numbers of single attached dwelling units and the numbers of apartments tend to be spatially correlated. For example, the simple correlation coefficient between the census tract magnitudes of single attached dwelling units and apartments was significant in Guelph, Peterborough, Brantford, Sault Ste. Marie, Sudbury and St. Catharines.

A third set of regression equations was developed by using the number of single detached dwelling units and the number of single attached plus apartments as the two independent variables. The properties of this set of regression equations are presented in Table 13. This table shows that the coefficient of determination varies from 0.90 in the London CMA to 0.99 in a number of census areas. A comparison of the magnitudes of b_1 in Tables 12 and 13 shows that there is little change in the magnitudes except for some of the smaller census areas. In Table 13 the b_1 magnitude varies from 1.439 in Thunder Bay to 1.767 in Sudbury and 1.743 in Kitchener. Twelve of the partial regression coefficient magnitudes presented in Table 13 fall between 1.5 and 1.6. Sufficient time was not available to determine if differences in household structure between the census areas accounted for the differences in b_1 magnitudes.

The magnitudes of b_2 in Table 13 vary from 0.895 in Windsor to 1.396 in Toronto. There is much more variability in this coefficient magnitude between census areas. Detailed reviews of household structure and the housing markets would be required in order to explain these differences.

The regression equations presented in Table 13 provide a useful and reliable means for estimating the spatial distributions of labour force given information on the dwelling unit composition of each census tract. The consistency and explanatory power of these equations suggest that work trip production studies be standardized in terms of dwelling unit composition. The dwelling unit composition of census tracts may be reasonably estimated for future time horizons from land use planning data.

6.3/ Travel Characteristics by Tenancy and Period of Residence

Table 14 summarizes the work trip linkage rates for the 15 Ontario urban areas for dwelling unit owners and renters. The entries in this table are based on linkages to work only within the census areas and do not include places of employment outside the census areas.

Table 13/ Census Tract Labour Force Prediction Equations Using Dwelling Unit Composition

Census Area	a	b_1	b_2	R^2
Guelph	32	1.572	1.176	0.96
Peterborough	9	1.537	0.907	0.98
Sarnia	-35	1.556	0.909	0.96
Brantford	43	1.503	1.005	0.99
Sault Ste. Marie	-26	1.561	1.301	0.99
Kingston	25	1.522	1.167	0.98
Thunder Bay	37	1.439	0.932	0.96
Oshawa	-30	1.525	1.118	0.96
Sudbury	-92	1.767	1.377	0.96
Kitchener	-95	1.743	1.335	0.98
Windsor	-69	1.577	0.895	0.97
London	149	1.521	1.048	0.90
St. Catharines	13	1.505	0.829	0.99
Hamilton	-22	1.555	1.098	0.92
Ottawa	112	1.664	1.168	0.94
Toronto	97	1.542	1.396	0.95

C.T. Labour Force = a + b_1 × C.T. Single Detached Dwelling Units

+ b_2 × C.T. Single Attached Dwelling Units and Apartments

Table 14/ Trip Generation Rates for Ontario Urban Areas Based on Tenancy

Census Area	Class	Linkages Produced	Number of Households	Trip Generation Rate
Guelph	Own	14 790	11 590	1.28
	Rent	6 720	6 625	1.01
Peterborough	Own	16 635	12 805	1.30
	Rent	5 610	6 005	0.93
Sarnia	Own	19 650	16 000	1.23
	Rent	6 675	6 635	1.01
Brantford	Own	19 755	16 550	1.19
	Rent	7 155	7 360	0.97
Sault Ste. Marie	Own	20 235	14 935	1.35
	Rent	6 510	6 165	1.06
Kingston	Own	16 125	15 520	1.04
	Rent	13 020	12 400	1.05
Thunder Bay	Own	30 585	23 575	1.30
	Rent	7 770	8 615	0.90
Oshawa	Own	25 965	22 725	1.14
	Rent	10 320	11 255	0.91
Sudbury	Own	30 480	22 860	1.34
	Rent	18 570	16 535	1.12
Kitchener	Own	54 015	39 905	1.35
	Rent	30 105	26 645	1.11
Windsor	Own	61 335	52 240	1.17
	Rent	20 145	21 930	0.92
London	Own	69 375	51 980	1.33
	Rent	37 545	35 155	1.07
St. Catharines	Own	77 805	63 895	1.22
	Rent	22 905	25 050	0.91
Hamilton	Own	114 510	93 140	1.23
	Rent	52 425	53 140	0.99
Ottawa	Own	120 780	85 595	1.41
	Rent	94 770	84 430	1.12

Table 14 shows that the trip generation rate per household is higher for owners than for renters with the exception of Kingston. The work trip linkage rate for owners varies from 1.04 in Kingston to 1.41 in Ottawa; the rate for renters varies from 0.90 in Thunder Bay to 1.12 in Sudbury and Ottawa. The higher average rates in owner occupied dwelling units are due to a variety of factors that might include larger household sizes with greater numbers in the labour force.

Table 15 summarizes the work trip generation rates for the 15 urban areas based on period of residence categories. The generation rates for the longer period of residence are uniformly higher than the 0 to 5 year category, except for Kingston. The work trip linkage rate varies from 1.11 in Windsor to 1.32 in Sault Ste. Marie for periods of residence six years and greater and from 0.99 in Oshawa to 1.25 in Sudbury for periods of residence of five years and less. The lower trip rates for the shorter periods of residency no doubt reflect the smaller household sizes of the recently established households.

Table 16 shows the mean trip lengths of households classified by tenancy type and of households classified by period of occupancy. Although this trip length information is unrelated to the generation rates, this section is the most convenient place to introduce this information for the different socio-economic groups. This table shows that the mean trip length to work of renters is significantly different from that of owners. The biggest difference is in Ottawa where many work trips from owner occupied dwelling units are from residences outside of the greenbelt.

Table 16 also shows that the average mean trip length to work from dwelling units that have been occupied for five years or less is in most cases longer than the average trip from households that have been occupied for six years or longer. This is reasonable since most of the new dwelling unit opportunities in urban areas are located on the periphery. The largest differences are in Windsor and Ottawa.

Table 15/ Trip Generation Rates for Ontario Urban Areas Based on Period of Residence

Census Area	Class	Linkages Produced	Number of Households	Trip Generation Rate
Guelph	0-5 yrs.	11 115	9 540	1.19
	6+	10 410	8 665	1.20
Peterborough	0-5	9 765	8 650	1.13
	6+	12 180	10 120	1.20
Sarnia	0-5	12 180	10 720	1.14
	6+	14 160	11 935	1.19
Brantford	0-5	12 000	10 685	1.10
	6+	14 910	13 230	1.12
Sault Ste. Marie	0-5	11 430	9 520	1.20
	6+	15 300	11 570	1.32
Kingston	0-5	17 175	14 585	1.18
	6+	11 970	10 380	1.15
Thunder Bay	0-5	15 900	14 180	1.12
	6+	22 410	18 080	1.24
Oshawa	0-5	17 010	17 110	0.99
	6+	19 275	16 820	1.15
Sudbury	0-5	26 305	21 075	1.25
	6+	23 715	18 340	1.29
Kitchener	0-5	46 080	37 995	1.21
	6+	38 040	29 985	1.30
Windsor	0-5	36 060	32 940	1.09
	6+	45 505	41 145	1.11
London	0-5	57 780	47 795	1.21
	6+	49 155	39 360	1.25
St. Catharines	0-5	43 065	39 115	1.10
	6+	57 660	49 765	1.16
Hamilton	0-5	82 275	73 943	1.11
	6+	84 675	72 320	1.17
Ottawa	0-5	122 655	98 595	1.24
	6+	92 895	71 425	1.30

Table 16/ Mean Trip Lengths for Socio-Economic Groups in Ontario Cities

Census Area	Mean Trip Length (km)			
	Owner	Renter	Occupancy <5 Years	Occupancy 6+ Years
Guelph	3.6	3.8	3.9	3.5
Peterborough	3.3	3.0	3.3	3.2
Sarnia	6.4	5.7	6.4	6.0
Brantford	4.2	3.8	4.2	4.1
Sault Ste. Marie	4.6	4.1	4.5	4.5
Kingston	5.8	5.2	5.6	5.4
Thunder Bay	5.4	4.6	5.2	5.3
Oshawa	4.1	3.7	3.9	4.1
Sudbury	8.6	7.9	8.7	8.2
Kitchener	4.7	4.7	4.8	4.5
Windsor	7.7	6.7	7.9	7.1
London	6.6	5.9	6.6	6.1
St. Catharines	6.3	6.1	6.1	6.0
Hamilton	8.3	6.9	7.9	7.7
Ottawa	8.4	6.4	7.9	7.1

6.4/ Travel Characteristics by Car Ownership Status

The most useful partitioned travel data available from the 1971 Census for transport planning purposes are the spatial interaction patterns by car ownership status. The ability to estimate transit captivity and noncaptivity from dwelling unit information is an extremely useful capability.

The regression analyses of work trip linkage formation rates described in Subsection 6.2 demonstrated that the best prediction equations could be developed in terms of the dwelling unit composition of census tracts where the dwelling unit composition is expressed in terms of the number of single detached units and the number of attached dwelling units plus apartments. The use of the number of attached dwelling units as a third independent variable was found to yield statistically nonsignificant partial regression coefficients for many of the census areas. An initial analysis of the trip productions by the two car ownership groups using three independent variables demonstrated similar behaviour.

Tables 17 and 18 illustrate the prediction equations developed to estimate the census tract amounts of captive and noncaptive amounts of labour force, respectively, along with the statistical properties of these equations for the 15 Ontario census areas. For the captive equations the partial regression coefficient of detached dwelling units varies from 0.156 in Kingston to 0.441 in Guelph; the t-statistics indicate that all these coefficients are significant at the 1% level. The partial regression coefficient of attached dwelling units varies from 0.331 in Oshawa to 0.639 in Guelph. Pooled regression equations are also shown for each census area size, class, and for all census areas combined; however, the F ratios are all significant, indicating that the pooled equations cannot be used instead of the individual equations. Inspection of the two partial regression coefficients shows that labour force captivity generation from single attached dwelling units and apartments is from two to three times greater than for single detached dwelling units.

The noncaptivity equations in Table 18 show that the partial regression coefficient of detached dwelling units varies from 0.800 in Oshawa to 1.267 in Ottawa, while the coefficient of attached dwelling units varies from 0.185 in Brantford to 0.542 in Sudbury. The attached dwelling unit partial regression coefficients are not significant for Guelph, Thunder Bay, Windsor and St. Catharines.

The noncaptive coefficients are two to three times larger for the detached dwelling units than the attached dwelling units. The table also illustrates that the pooled regression equations cannot be used in place of the equations for individual census areas.

Table 19 summarizes the mean home to work transport network distances for non car owners and car owners for the 15 Ontario census areas. This table shows that there are some large differences between the mean trip lengths for the two groups in some areas. The largest differences are in Hamilton, Sudbury, Ottawa, Sarnia and Windsor.

Table 17/ Census Zone Captivity Prediction Equations Using a Modified Dwelling Unit Composition

Census Area	α	b_1 (t_1)	b_2 (t_2)	R^2	F Ratio	DF
1. Guelph	-177	0.441 (5.38)	0.639 (5.72)	0.87		
2. Peterborough	13	0.299 (4.90)	0.421 (5.06)	0.81		
3. Sarnia	16	0.209 (11.10)	0.474 (13.14)	0.97		
4. Brantford	35	0.213 (5.03)	0.593 (9.41)	0.91		
5. Sault Ste. Marie	-104	0.287 (7.86)	0.705 (10.41)	0.95		
6. Kingston	-30	0.156 (5.32)	0.621 (14.76)	0.95		
GROUP 1–6	4	0.224 (12.62)	0.569 (21.06)	0.88		
CHOW TEST 1–6					3.121	15, 72
7. Thunder Bay	10	0.273 (6.88)	0.555 (7.19)	0.94		
8. Oshawa	-133	0.340 (7.76)	0.331 (5.16)	0.87		
9. Sudbury	-40	0.275 (5.13)	0.562 (9.16)	0.86		
10. Kitchener	50	0.226 (2.80)	0.661 (7.62)	0.78		
GROUP 7–10	-94	0.300 (9.49)	0.605 (15.55)	0.85		
CHOW TEST 7–10					4.840	9, 70
11. Windsor	-98	0.269 (6.51)	0.525 (9.95)	0.85		
12. London	17	0.255 (8.18)	0.521 (16.04)	0.94		
13. St. Catharines	9	0.178 (6.32)	0.537 (8.15)	0.93		
14. Hamilton	-38	0.265 (5.83)	0.536 (7.25)	0.71		
15. Ottawa	54	0.206 (5.40)	0.606 (18.91)	0.89		
GROUP 11–15	-13	0.219 (12.80)	0.589 (29.07)	0.86		
CHOW TEST 11–15					1.585	12, 156
ALL CASES	-5	0.224 (19.82)	0.586 (42.44)	0.88		
CHOW TEST – ALL CASES					1.552	42, 298

ZONE CAPTIVES = $a + b_1 \times$ Zone Single Detached Dwelling Units $+ b_2 \times$ Zone Single Attached Dwelling Units and Apartments

6.5/ Generalizations about Trip Generation

The trip generation analyses described in Section 6 clearly demonstrate that high quality prediction equations for estimating the labour force composition of census tracts may be developed in terms of the dwelling unit composition of census tracts. The important feature of these prediction equations is that the partial regression coefficients are consistent across all the Ontario census areas. Any differences are due to variations in the labour force participation rates between census areas. High quality prediction equations have also been developed in terms of dwelling unit composition for estimating labour force transit captivity and noncaptivity.

These findings suggest that work trip generation rates should be standardized in terms of dwelling unit composition for all urban transport studies. The dwelling unit composition of census tracts and traffic analysis zones is perhaps the highest quality land use data available in most municipalities. More important, it is a variable that is influenced strongly by municipal planning policies and may be reliably estimated for future time horizons. The dwelling unit is the focus of municipal assessment records; therefore,the opportunities for developing more sophisticated dwelling unit classification schemes and relating travel behaviour to dwelling unit classifications are enormous. Assessment records also contain information on household structure, which would allow reliable correlations to be established between dwelling unit type and household structure variables such as labour force composition.

Table 18/ Census Zone Non-Captivity Prediction Equations Using a Modified Dwelling Unit Composition

Census Area	α	b_1 (t_1)	b_2 (t_2)	R^2	F Ratio	DF
1. Guelph	143	0.838 (6.75)	0.319 (1.89)	0.83		
2. Peterborough	-12	0.989 (13.72)	0.262 (2.65)	0.94		
3. Sarnia	-79	1.073 (21.96)	0.328 (3.51)	0.98		
4. Brantford	18	0.953 (22.13)	0.185 (2.89)	0.98		
5. Sault Ste. Marie	86	1.023 (13.16)	0.353 (2.45)	0.94		
6. Kingston	32	1.185 (27.87)	0.247 (4.04)	0.98		
GROUP 1–6	-22	1.054 (37.72)	0.302 (7.09)	0.95		
CHOW TEST 1–6					3.364	15, 72
7. Thunder Bay	-165	1.068 (13.89)	-0.008 (-0.05)	0.95		
8. Oshawa	146	0.800 (9.49)	0.415 (3.35)	0.88		
9. Sudbury	-94	1.076 (12.58)	0.542 (5.53)	0.91		
10. Kitchener	31	1.167 (10.42)	0.358 (2.97)	0.86		
GROUP 7–10	-69	0.982 (18.58)	0.558 (8.57)	0.87		
CHOW TEST 7–10					6.278	9, 70
11. Windsor	67	0.938 (15.37)	0.129 (1.65)	0.90		
12. London	130	1.039 (12.69)	0.320 (3.74)	0.89		
13. St. Catharines	118	1.022 (17.87)	0.036 (0.27)	0.96		
14. Hamilton	-281	1.046 (19.19)	0.393 (4.43)	0.92		
15. Ottawa	-53	1.267 (16.06)	0.365 (5.52)	0.86		
GROUP 11–15	-7	1.017 (31.78)	0.384 (10.17)	0.87		
CHOW TEST 11–15					6.080	12, 156
ALL CASES	-20	1.020 (49.64)	0.393 (15.64)	0.90		
CHOW TEST – ALL CASES					3.779	42, 298

ZONE NON-CAPTIVES = $a + b_1 \times$ Zone Single Detached Dwelling Units $+ b_2 \times$ Zone Single Attached Dwelling Units and Apartments

Table 19/ Mean Trip Lengths for Non-Car Owners and Car Owners in Ontario Census Areas

Census Area	Mean Trip Length (km) Non-Car Owner	Car Owner
Guelph	3.4	3.9
Peterborough	3.0	3.4
Sarnia	4.5	6.7
Brantford	3.5	4.4
Sault Ste. Marie	3.5	4.9
Kingston	4.3	6.0
Thunder Bay	4.6	5.6
Oshawa	3.9	4.1
Sudbury	6.4	9.2
Kitchener	4.1	5.1
Windsor	6.0	8.0
London	5.1	6.9
St. Catharines	4.7	6.8
Hamilton	5.8	8.8
Ottawa	6.1	8.3

7/ Empirical Analyses of Spatial Interaction

The cell entries in observed trip matrices reflect two types of effect:
(1) the rate of interaction between a pair of zones; and
(2) the labour force and employment magnitudes of the zones.

To better understand the spatial interaction patterns, it is essential to extract from the home to work linkage matrices the pure spatial interaction effects that are independent of zone size effects. It is these rates of spatial interaction that trip distribution models attempt to explain in terms of the travel deterrence function.

This section describes some empirical studies of the journey-to-work matrices obtained in the 1971 census for the 15 Ontario census areas. The two basic steps of the analysis strategy are:
(i) bi-proportional balancing of the trip matrices; and
(ii) the clustering of zones by similarities in destination characteristics.

The bi-proportional balancing techniques transform the home to work linkages matrices so that the entries in each row and in each column sum to a constant magnitude. The effect of this operation is to produce matrices in which the entries represent the interaction magnitudes that would occur between zones of equal size, or the pure interaction effects between home and work zones. The clustering technique compares the destination vectors for each origin zone and groups zones together that have the most similar destination vectors. The clustering technique groups all zones into a single cluster using a step by step approach.

7.1/ Bi-Proportional Matrix Balancing

It was suggested that bi-proportional balancing techniques may be used to remove the effects of zone size from trip matrices, leaving only the information on the rate of spatial interaction between zones. If this is not done, zones with large amounts of employment will dominate any analyses and may mask important interaction effects between smaller zones. Bi-proportional matrix balancing techniques have been developed for use in a number of fields with much of the original work conducted in connection with input-output tables [2]. The Furness method [3] of origin-destination table development is a closely related procedure.

The entries in bi-proportionally adjusted trip matrices represent the interactions that would occur if all the zonal productions and attractions in an urban area were equal. The row/column constant used in matrix balancing can have any arbitrary magnitude. For example, if it were equal to 1.0, then each cell entry could be interpreted as the probability of interaction between a known origin zone and a particular destination zone when all destination attractions have the same magnitude. Normalized trip matrices, therefore, reflect only spatial separation effects rather than both spatial separation and structural effects.

The calculation of bi-proportional matrices is straightforward. Trip matrix row entries are each scaled so that the row totals are all equal to the arbitrary constant. The columns of this new matrix are then summed and scaled so that the column totals are all equal to the arbitrary constant. This process is repeated until convergence is obtained with all row and column totals being equal to the arbitrary constant.

It is useful to examine the properties of the transformed matrices in relation to the original trip matrix. The bi-proportional balancing techniques preserve the interaction rate structure of the original matrix, but the cell entries in any row or column are not all adjusted by a factor with a constant magnitude. However, it has been shown [2] that the products of the final row and column adjustment factor magnitudes have unique magnitudes for each cell. That is, the interaction magnitudes calculated by the bi-proportional balancing techniques are unique.

It is also useful to examine the changes in the trip length structure that might be expected from the transformation of the original trip matrix. The mean trip length of any origin or destination zone is calculated simply from the weighted sum of the trip distances to each of the other zones in the system. It was noted previously that the bi-proportionalized matric retains the interaction rate structure but not the interaction magnitude structure of the original matrix retains the interaction rate structure but not or destination zone for the bi-proportionalized matrix may be expected to change, since the row or column weight composition used will have altered. Rows that do not have strong concentrations of destinations at one or two zones may be expected to exhibit little change in the zones mean trip length. On the other hand, zones with strong concentrations of destinations in a few zones might be expected to exhibit significant shifts in the mean trip length. The weights between zones with small trip end magnitudes will increase and exert a greater influence on the zone mean trip lengths of the bi-proportionalized matrices. The bi-proportional adjustment process will produce more uniform zone mean trip lengths than originally observed.

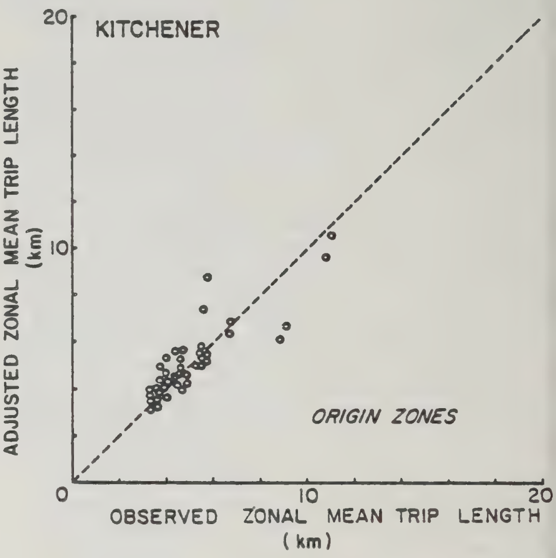
Table 20 summarizes the area-wide trip lengths for both the original trip matrix and the balanced matrix for the 15 Ontario census areas. This table shows that in most cases the area-wide mean trip lengths of the balanced matrices are shorter than the observed. The largest shifts occur in those census areas with unusual trip linkage characteristics such as Sarnia, Kingston, Sudbury, Hamilton and Ottawa. In these areas a few employment zones tend to dominate the commuting structure.

Figures 19 to 21 provide more detailed comparisons of the adjusted and observed trip matrices in which the adjusted mean trip length for each zone is plotted against the observed mean trip length. Figures 19 and 20 provide comparison for the origin zones in four census areas. Figure 21 shows comparisons for the destination zones of Hamilton and Ottawa.

Figures 19 and 20 illustrate the tendency for shorter trip lengths to become longer and the longer trip lengths shorter. The largest shifts in the mean trip lengths are associated with zones located on the periphery or the smaller-sized zones. Figure 21 illustrates that the mean trip lengths for the destination zones tend to be reduced by the balancing procedure. It should be remembered that the aim of the bi-proportional balancing procedure is not to preserve the trip length but the interaction structure of a community.

Table 20/ Comparison of Area-Wide Observed and Balanced Trip Matrix Mean Trip Lengths

	Observed Mean Trip Length (km)	Balanced Mean Trip Length (km)
Guelph	3.76	3.77
Peterborough	3.30	3.36
Sarnia	6.33	5.44
Brantford	4.20	3.71
Sault Ste. Marie	4.59	4.55
Kingston	5.68	4.32
Thunder Bay	5.37	5.65
Oshawa	4.10	4.10
Sudbury	8.63	7.09
Kitchener	4.80	5.03
Windsor	7.50	7.31
London	6.57	6.70
St. Catharines	6.50	6.23
Hamilton	8.05	7.10
Ottawa	7.63	6.88



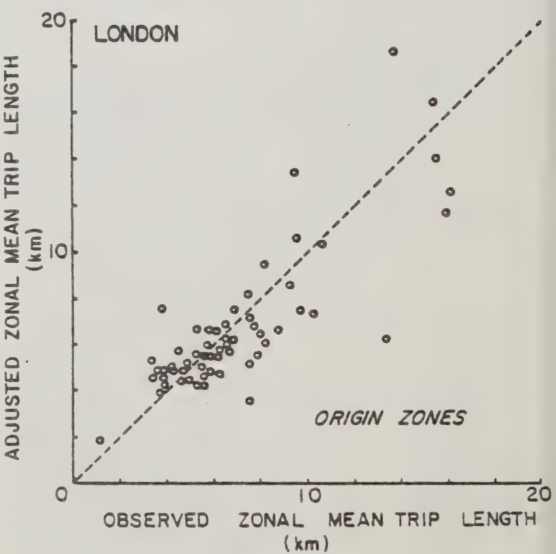
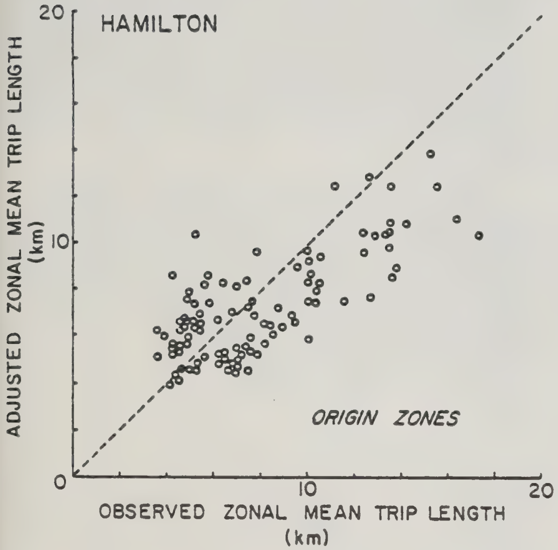
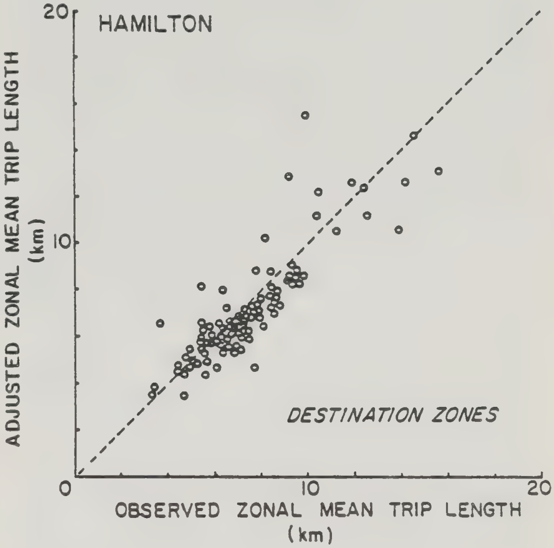
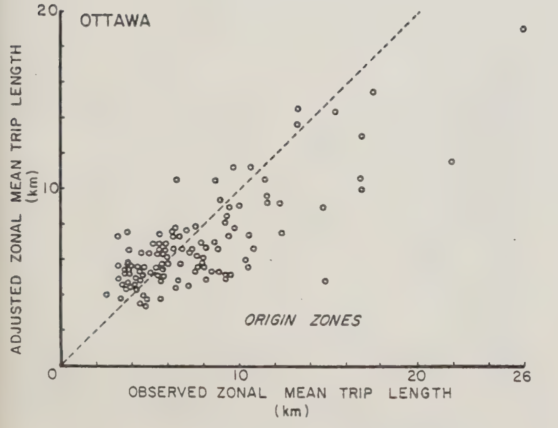


Figure 19/ Adjusted vs Observed Origin Zone Mean Trip Lengths for Kitchener and London







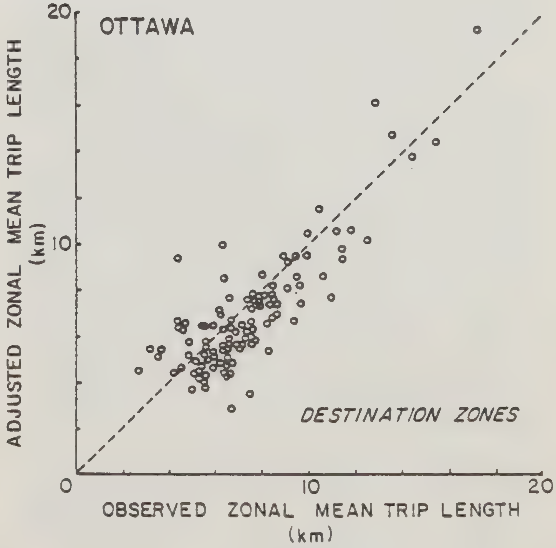


Figure 20/ Adjusted vs Observed Origin Mean Trip Lengths for Hamilton and Ottawa

Figure 21/ Adjusted vs Observed Destination Zone Mean Trip Lengths for Hamilton and Ottawa

7.2/ Clustering Techniques

The clustering technique used in this investigation is a procedure known as Ward's method, which is a hierarchical agglomerative technique [4]. This technique begins with the n census tracts in a census area and groups them into a series of clusters with aggregation continuing until all of the census tracts are grouped into one cluster. With this technique the analyst is not constrained to some arbitrary number of clusters.

The clustering technique uses a measure of Euclidean distance between the destination-zone similarities of two census tracts that is defined in the following way:

$$d_{ij} = \left[\sum_{k=1}^{n} \frac{(T^*_{ik} - T^*_{jk})^2}{n} \right] \tag{3}$$

where

d_{ij} = the Euclidean distance between the destination characteristics of the two census tracts i and j

k = the destination zone under consideration for a particular origin zone (i or j)

T^*_{ik} = the normalized number of trips between origin zone i and destination zone k

n = the total number of census tracts

That is, d_{ij} is a measure of the similarity in the destination vectors of the two particular origin zones being considered for a potential merger. If the destination vectors were identical, then $d_{ij} = 0$. The lower the magnitude of d_{ij}, the greater the similarity between the two census tracts.

A detailed description of the clustering procedure is provided in Appendix A. It is pointed out in Appendix A that the distance measure calculated by Equation 3 may be interpreted as a measure of the variation between the destination vectors of two census tracts. As more census tracts are included in the cluster, the distance value calculated by Equation 3 will tend to increase and the clustering procedure groups census tracts so as to minimize the increase in the error across the entire set of census tracts. One would expect that the first stages in the clustering hierarchy would consist of a number of pairwise census tract clusters, each with a small increase in the error measure. Large increases in the aggregate sum of squared differences might be expected as clusters continue to be grouped to form eventually a single cluster.

7.3/ Dendrograms for Ontario Census Areas

The results of cluster analyses may be plotted conveniently in the form of a dendrogram, which traces the clustering hierarchy and identifies the magnitudes of the error at which merging occurs. The dendrograms for the Guelph and Brantford CAs are illustrated in Figure 22 and the increases in the error with decreasing numbers of clusters are illustrated in Figure 23.

The dendrogram for Guelph in Figure 22 shows that the first pair of census tracts to cluster are 4 and 13, followed by 7 and 8, 5 and 12, and 9 and 11. As the clustering process continues, 10 clusters with the already clustered 9 and 11. This dendrogram shows that clustering continues without major jumps in the error sum of squares. With three clusters the census tract composition of each cluster is:

I: 1,100, 3, 4, 13, 5 and 12

II: 2, 9, 11 and 10

III: 6, 7 and 8

The clustering process in the Brantford region has slightly different characteristics than that for Guelph. The initial stages of the clustering are similar, but the error sum of squares increases sharply during the final stages of the clustering process. The dendrogram for Brantford suggests the following census tract composition for three clusters:

I: 1, 8, 10, 9, 2, 6 and 7

II: 3, 4, 13, 5, 14, 12, 100 and 110

III: 11

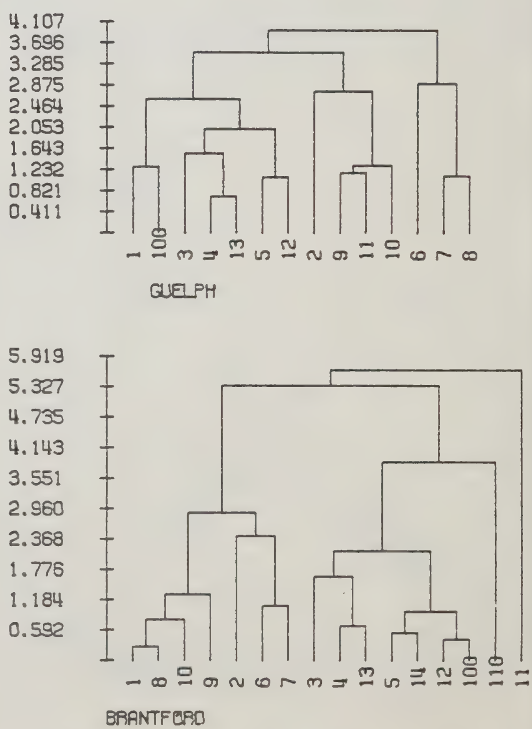


Figure 22/ Dendrograms for Guelph and Brantford Census Areas

It is interesting to note that census tract 11 in Brantford is almost exclusively an industrial zone and that it clusters with the remaining census tracts only at the last stage. It should also be noted from Figures 22 and 23 that the increase in the error sum of squares is large when clusters I and II are merged. Also, a similarly large increase occurs when census tract 110 is merged with the other members of cluster II in the previous stage. Census tract 110 embraces the Paris area, which is part of the Brantford census area, and its forced merger with the census tracts in the north-western part of Brantford produces the relatively large increase in the error sum of squares at the pre-penultimate state of clustering.

The dendrograms for each Ontario census area are contained in Appendix B with their interpretations. The characteristics of the Guelph and Brantford dendrograms sufficiently highlight the general properties of the dendrograms contained in Appendix B. Some of the effects highlighted above are more extreme in the larger census area than in the two smaller census areas. For example, the Kitchener census area consists of the three major municipalities of Waterloo, Kitchener and Cambridge, which embrace relatively self-contained commutersheds. When these areas are forced together, large increases in the error sum of squares occur. The ordinates of all of the dendrograms presented in Appendix B are plotted in the same scale and differences in the clustering properties of each census area are detected easily.

The residential zone clustering properties of four census areas (Kitchener, London, Hamilton and Ottawa) are described in detail in the following paragraphs. These areas have been selected to highlight the clustering characteristics observed for areas with different spatial characteristics and of different sizes.

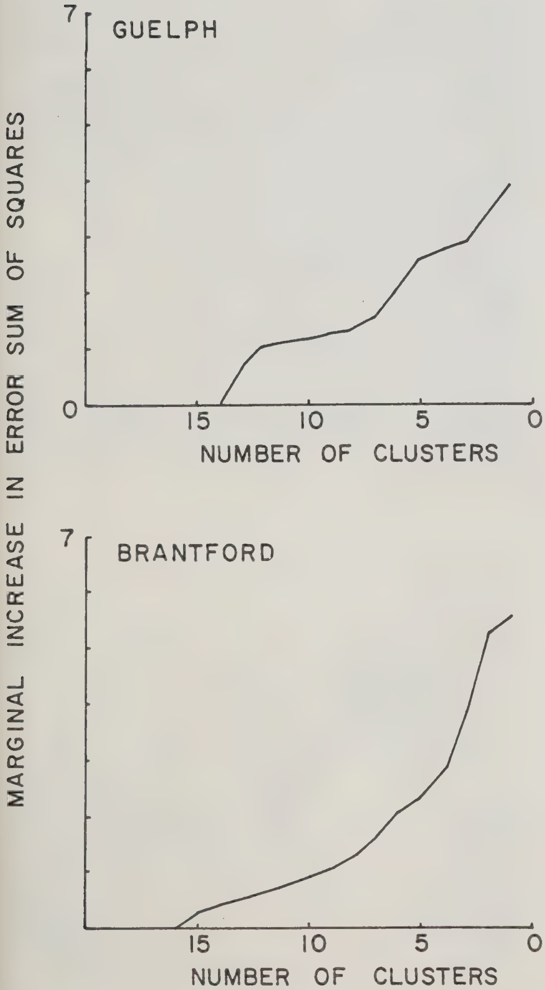


Figure 23/ Increases in Error with Decreasing Numbers of Census Tracts

7.4/ Detailed Analyses of Four Census Areas

The Kitchener CMA consists of the Cities of Kitchener, Waterloo and Cambridge. A review of the Kitchener CMA dendrogram in Appendix B shows that a number of very distinct clusters emerge from the analysis of the area and that the error measure increases sharply as clusters continue to be merged after about six clusters. The upper part of Figure 24 illustrates the boundaries at the 10 cluster level for the Kitchener CMA. One cluster is not shown since it consists of a group of rural zones with small towns and villages in the southern part of the census area. At the 10 cluster level there are two clusters in Waterloo, four in Kitchener, one embracing the village of Doon and two in Cambridge. This diagram also illustrates the mergings that occur at the 9 and 8 cluster levels that involve further groupings of the Kitchener clusters.

The lower part of Figure 24 shows the cluster boundaries at the 5 cluster level. It is important to note that these boundaries are almost coincident with the boundaries of the five municipalities that existed in 1971. Transition from the 8 cluster to 5 cluster level occurred largely through mergings within the City of Kitchener.

To improve our understanding of commuting patterns it is important to isolate the factors that determine commuting subregions of the type

illustrated in Figure 24. These subregions really reflect two dimensions in space since they bound census tracts with similar destination zones. As a result, any analysis of potential determinants should attempt to reflect this two-dimensional character.

Factors that seem to influence the development of commuting linkages include the timing of development, socio-economic factors, topographic and man-made barriers, and municipal boundaries. The only systematic data available that might be used for these analyses are the variables describing households collected in the household census. Household end variables that might capture this two-dimensional structure of commuting must be used.

The two most important variables available from the census tract bulletins are occupation and industry group. Cluster analyses were performed for the four CMA analyzed in detail in this section using the following characteristics of census tracts in two separate analyses:

1/ a vector of the proportions of the residential labour force in 16 occupation groups;

2/ a vector of the proportions of the residential labour force employed in 8 industry groups.

Cluster analyses were also performed on variables such as period of residence, dwelling unit type and automobile ownership.

Figures 25 and 26 illustrate the hierarchical structures of the Kitchener CMA yielded by these analyses based on occupation type and industry sector, respectively. These two diagrams show that there is no unambiguous delineation of the commuting subregions by these cluster analyses of home-end work-related variables. Figure 27 illustrates the hierarchical structure yielded by the cluster analysis of the period of residence characteristics of residential census tracts. Again, there is no consistent delineation of commuting subregions. The strongest common elements of the hierarchical structures illustrated in the four diagrams are the clustering of census tracts in Waterloo and those in the eastern half of Cambridge. The principal determinants of these two areas are growth at the University of Waterloo during 1966–1971 and parallel development of residential areas in Waterloo, and the growth in manufacturing in Cambridge. It seems clear that while these cluster analyses of residential variables provide some insight into the commuting structure of the Kitchener CMA, more elaborate analyses would be

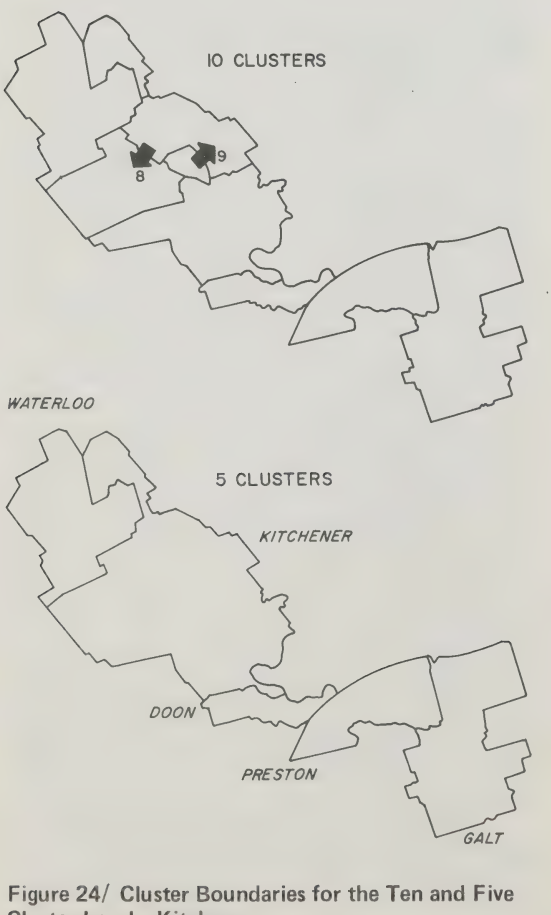


Figure 24/ Cluster Boundaries for the Ten and Five Cluster Levels, Kitchener

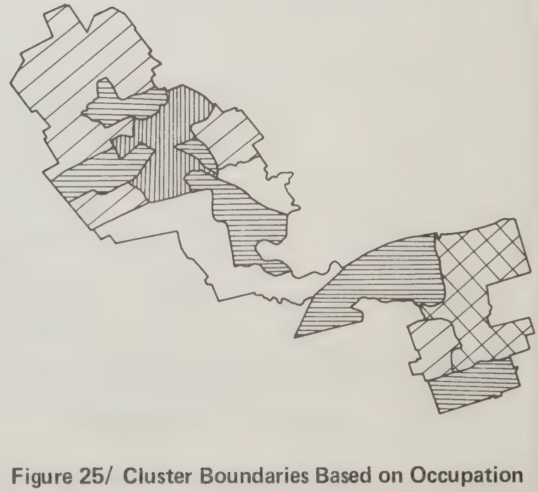


Figure 25/ Cluster Boundaries Based on Occupation Type, Kitchener

required to capture simultaneously the influences of several factors such as timing of development, industry type and socio-economic characteristics.

Figure 28 illustrates the commuting structure of the London CMA at the 10 and 5 cluster levels. The subregions delineated in Figure 28 are generally contiguous, although there are some census tracts that cluster together but which are separated spatially. For example, the hatched subregions in Figure 28 belong to the same commuting cluster in spite of the fact that they are on opposite sides of the area. These are residential areas with strong linkages to CBD jobs. The other clusters are radial in nature and reflect the dominance in London of the CBD as an employment centre. These radial clusters separate largely on the basis of secondary linkages with other employment centres located in the individual sectors.

Figure 26/ Cluster Boundaries Based on Industry Sector, Kitchener

Figure 27/ Cluster Boundaries Based on Period of Residence, Kitchener

Figure 28/ Cluster Boundaries for the Ten and Five Cluster Levels, London

Figures 29 and 30 illustrate the spatial structure of London when clusters were formed on the basis of the occupation composition and industry sector composition of residential census tracts, respectively. The occupation group clusters are much more fragmented than the commuting clusters; however, the industry sector clusters reflect the radial structure of the commuting clusters to some extent. Once again the generation of clusters on the basis of household end variables do not provide reasonable delineations of the commuting subregions illustrated in Figure 28.

Figure 31 shows the commuting subregions for the Hamilton CMA at the 5 cluster level. There are three clusters in the Hamilton area; on Hamilton Mountain, the east end residential areas and the west end residential areas. There are two clusters in the Burlington area with the southern-most cluster separating on the basis of its commuting linkages with Hamilton. Merging of the Hamilton and Burlington clusters occurs at a relatively high error sum of squares level.

Figures 32 and 33 illustrate the clusters formed on the basis of occupation group and industry sector composition, respectively. The clusters formed using these variables are more complex than for the smaller census areas of Kitchener and London. In Hamilton the occupation clusters reflect to some extent the commuting structure illustrated in Figure 31, but the clusters are very fragmented.

Census tracts within the Hamilton CMA have also been clustered on the basis of automobile ownership. Figure 34 illustrates the structure at the 3 cluster level. These clusters delineate the inner suburbs of Hamilton with the outer suburbs of Hamilton clustering with the central areas of Burlington and the outer areas of Burlington forming the third cluster. These clusters based on automobile ownership provide additional insights into the commuting patterns of the Hamilton CMA, but again, the household based variables are not strong enough to reflect the two-dimensional character of the commuting patterns.

Figure 35 illustrates the commuting structure of the Ottawa CMA at the 8 cluster level. These clusters reflect fairly clearly a number of well-organized factors in the Ottawa area such as the presence of the Quebec-Ontario border and the Ottawa River, as well as the inner city-outer suburbs dichotomy.

Figures 36 and 37 illustrate the structure of the Ottawa CMA when residential census tracts are clustered on the basis of occupation group and industry sector composition, respectively. The clusters formed in this way are very fragmented and do not reflect in any unambiguous way the commuting subregions illustrated in Figure 35. Figure 38 illustrates the structure of the Ottawa CMA revealed by a cluster analysis based on the dwelling unit composition of census tracts. This variable on its own does not provide any additional insights into the commuting structure of the Ottawa CMA. In combination with other variables, however, it would no doubt help to delineate the commuting subregions.

Figure 29/ Cluster Boundaries Based on Occupation Type, London

Figure 30/ Cluster Boundaries Based on Industry Sector, London

7.5/ Some Generalizations about Commuting Patterns

The cluster analyses and some general knowledge about the other Ontario census areas allow several generalizations to be made about the determinants of commuting patterns. Table 21 summarizes these determinants for the 15 census areas. Seven of the census areas as defined by Statistics Canada embrace two or more separate municipalities. In almost all cases, major cluster boundaries were coincident with the municipal boundaries indicating that commuting patterns tend to be self-contained within municipalities. Thunder Bay, Kitchener, St. Catharines and Hamilton provide good examples of this type of effect.

Topographic features such as rivers also have an important influence on commuting patterns. The best examples are provided by London, Hamilton and Ottawa. Ottawa is a special case since the Ottawa River is also a provincial boundary.

One of the most important influences on commuting patterns is the timing of development where residential areas and employment zones that expanded during the same time period tend to have strong linkages. The best example is Guelph where the residential areas and employment zones that grew rapidly during 1966–1971 were all located on the periphery. The commuting patterns that emerged as a consequence of this phase of development have a distinctly different character to the centrally focussed patterns that developed earlier.

Perhaps the strongest influences on the commuting structure of a community are the socio-economic factors that influence dwelling unit location possibilities. The importance of socio-economic factors is analyzed in more detail in subsequent sections of this report.

The final factor identified in Table 21 is the impact that specific employment concentrations have on commuting patterns. Communities with large concentrations of employment in one area either because of one industry or a collection of employers in that area develop particular types of commuting patterns that must be recognized in any modelling attempts. Sarnia, Sault Ste. Marie, Oshawa, Sudbury, Windsor and Hamilton are examples of areas with strong concentrations of industries in particular areas. The commuting patterns of these areas tend to dominate the home to work linkages patterns. London and Ottawa are examples in which employment in the tertiary sector dominates the economic base of the community and this employment type tends to be concentrated in the CBD.

Table 21/ Major Determinants of Residential Zone Commuting Clusters

Census Area	Multi-Community Composition of Census Area	Topographic Features	Timing of Development	Socio-Economic Factors	Specific Employment Concentrations
Guelph			X	X	
Peterborough			X	X	
Sarnia			X	X	X
Brantford	X	X	X	X	
Sault Ste. Marie				X	X
Kingston		X	X	X	
Thunder Bay	X			X	
Oshawa	X		X	X	X
Sudbury				X	X
Kitchener	X		X	X	
Windsor			X	X	X
London		X	X	X	X
St. Catherines	X			X	
Hamilton	X	X	X	X	X
Ottawa	X	X	X	X	X

The five determinants identified above and the associated clusters described previously and in Appendix B provide a basis for identifying calibration subregions within each of the census areas. The particular subregions used in the calibration of multi-parameter gravity models are described later in this report.

The heterogeneity in commuting patterns in each of the census areas is evident in Table 22 in which the maximum value of error sum of squares is reported for each census area with the number of spatial linkages in each area. Clearly this magnitude will increase with increasing size of urban area but deviations from the general trend may be clearly identified. For example, in the smaller urban areas the higher than average maximum error sum of squares for Sarnia, Sault Ste. Marie and Kingston may be noted. In the next size group Thunder Bay, Oshawa and Kitchener are examples of census areas in which commuting heterogeneity leads to sharp increases in the error sums of squares in the final stages of the clustering process. Finally, in the three largest census areas analyzed the error sums of squares also increased sharply in the final clustering stages primarily because residential zones from separate municipalities were being forced together.

Table 22/ Maximum Value of Error Sum of Squares for Ontario Census Areas

Census Area	Number of Home-Work Linkages	Error Sum of Squares
Guelph	19 990	3.9
Peterborough	20 400	5.3
Sarnia	23 760	6.3
Brantford	24 195	5.6
Sault Ste. Marie	24 740	6.5
Kingston	26 770	6.8
Thunder Bay	33 130	15.6
Oshawa	33 740	10.9
Sudbury	45 580	8.5
Kitchener	81 095	22.0
Windsor	74 455	10.3
London	97 195	8.1
St. Catharines	91 815	21.4
Hamilton	152 265	20.0
Ottawa	201 015	19.2

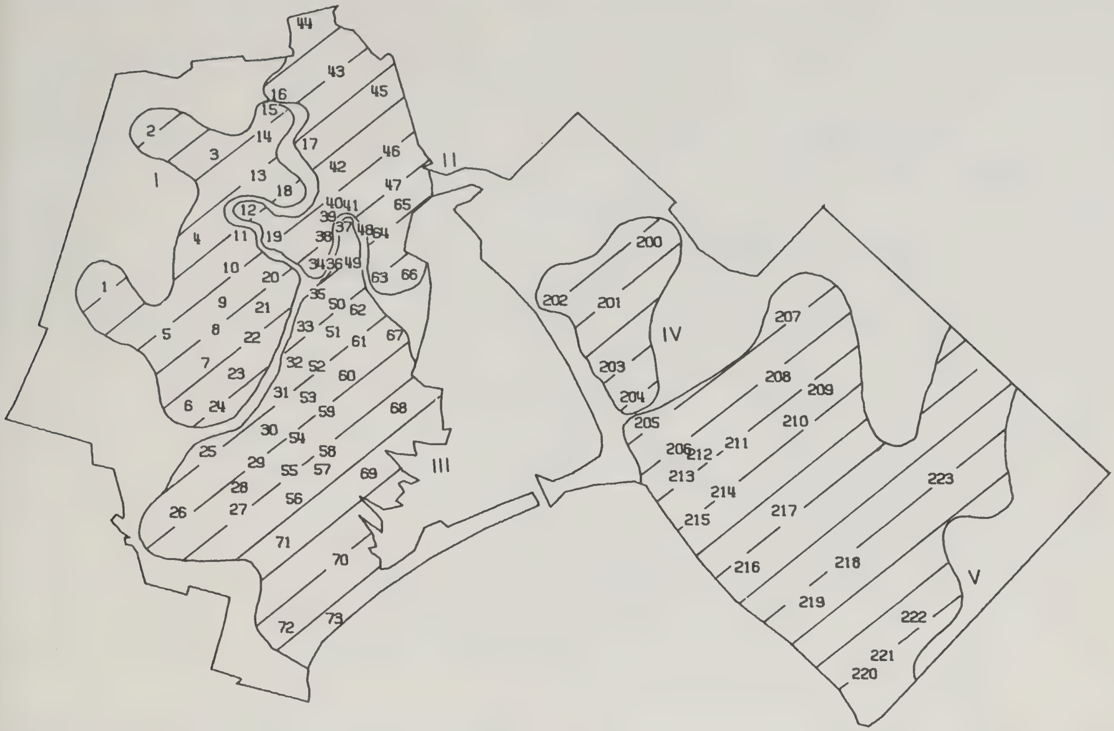


Figure 31/ Commuting Structure of Hamilton

Figure 32/ Cluster Boundaries Based on Occupation Type, Hamilton

Figure 33/ Cluster Boundaries Based on Industry Group, Hamilton

Figure 34/ Cluster Boundaries Based on Auto Ownership, Hamilton

Figure 35/ Commuting Structure of the Ottawa CMA

Figure 36/ Cluster Boundaries Based on Occupation Type, Ottawa

Figure 37/ Cluster Boundaries Based on Industry Sector, Ottawa

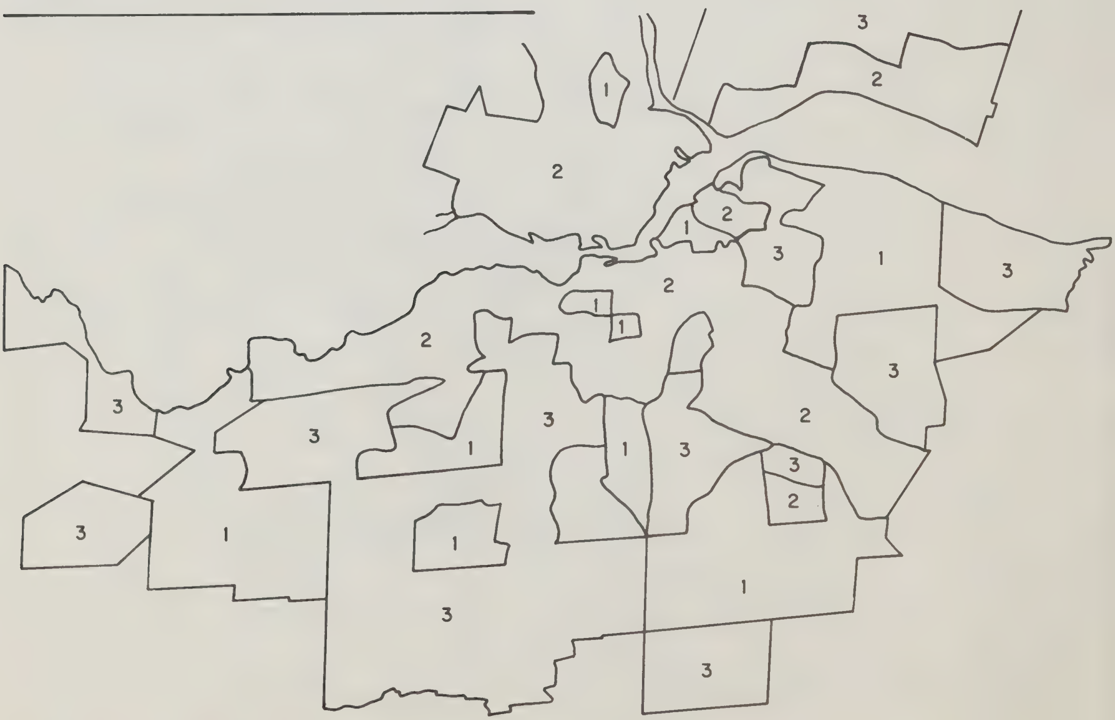


Figure 38/ Cluster Boundaries Based on Dwelling Unit Type, Ottawa

8/ Aggregate Trip Distribution Models

Conventional aggregate trip distribution models of the gravity type are estimated in this section. The calibration procedure used and the goodness of fit properties of the calibrated models are described. The explanatory qualities of production-constrained, attraction-constrained and doubly-constrained versions of the gravity model using distance as the inter-zonal cost measure are first examined. Changes in the calibration qualities are then examined when travel times and a combination of travel time and distance are used respectively as the cost measures. A variety of stratified distribution models are examined in Section 9.

8.1/ Basic Gravity Model Structure

The trip distribution model that is calibrated to the 1971 journey-to-work data for all census areas is a doubly-constrained version of the gravity model with a negative exponential deterrence function of the following form:

$$T^*_{ij} = A_i B_j O_i D_j \exp(-\beta_s c_{ij}) \tag{4}$$

where T^*_{ij} = the model estimated trip interchanges between zones i and j
O_i = trip productions of zone i
D_j = trip attractions of zone j
β_s = travel deterrence function parameter specific to all the origin zones in a given calibration subregion s
c_{ij} = the travel *costs* (distance, time or some combination) for a trip between zones i and j

and

$$A_i = \left[\sum_j B_j D_j \exp(-\beta_s c_{ij})\right]^{-1} \tag{5}$$

$$B_j = \left[\sum_i A_i O_i \exp(-\beta_s c_{ij})\right]^{-1} \tag{6}$$

two balancing factors which ensure that the trip end constraints

$$O_i = \sum_j T_{ij} \tag{7}$$

and

$$D_j = \sum_i T_{ij} \tag{8}$$

are satisfied.

Production-constrained and attraction-constrained versions of the gravity model have been estimated for eight Ontario census areas as well. The basic form of the production-constraint form of the gravity model estimated in this study is:

$$T^*_{ij} = A_i O_i D_j \exp(-\beta c_{ij}) \tag{9}$$

$$\text{where } A_i = \left[\sum_j D_j \exp(-\beta c_{ij})\right]^{-1} \tag{10}$$

while the basic form of the attraction-constrained gravity model is:

$$T^*_{ij} = B_j O_i D_j \exp(-\beta c_{ij}) \tag{11}$$

$$\text{where } B_j = \left[\sum_i O_i \exp(-\beta c_{ij})\right]^{-1} \tag{12}$$

8.2/ Model Calibration Technique

For the doubly-constrained model the magnitudes of the balancing factors A_i and B_j are estimated by assuming initially that all B_j are equal to 1, and calculating all of the A_i for a specific set of β_s values. These A_i magnitudes are then substituted into Equation 5 to calculate a new set of B_j magnitudes. This process continues in an iterative manner until the product magnitude A_iB_j converges for all i = j pairs. Convergence is assumed when the difference between the products A_iB_j for two successive iterations is less than 1.0×10^{-7}.

The balancing factors A_i and B_j may be set to 1.0 to produce, respectively, the production-constrained and attraction-constrained versions of the gravity model.

The gravity model is calibrated by establishing the magnitudes of β_s that minimize the sum of the absolute differences between the observed and model-estimated ordinates of the trip length frequency distributions. These frequency distributions are specified in 1 km class intervals. Where more than one calibration subregion is used, the observed and simulated frequency distributions are calculated separately for each β_s using only trips originating from those zones to which each parameter is specific.

The model parameters are estimated using a golden section search technique[5]. With this procedure an initial estimate is made of the maximum and minimum limits for the parameter, β_{01} and β_{02}. Two new values are then chosen such that:

$$\beta_{11} = 0.382(\beta_{02} - \beta_{01}) + \beta_{01} \tag{13}$$

$$\beta_{12} = 0.618(\beta_{02} - \beta_{01}) + \beta_{02} \tag{14}$$

The magnitudes 0.382 and 0.618 are derived from $(\Gamma-1)/\Gamma$ and $1/\Gamma$ where $\Gamma = 1.618$ the asymptotic interval reduction factor of the Fibonacci search method. The model estimated trip length frequency distributions are calculated using β_{11} and β_{12}. Whichever magnitude gives the worst fit is chosen as the new limit of search on that side of the optimum. For example, if β_{12} were the worst the next two parameter magnitudes would be:

$$\beta_{21} = 0.382(\beta_{12} - \beta_{01}) + \beta_{01} \tag{15}$$

$$\beta_{22} = 0.618(\beta_{12} - \beta_{01}) + \beta_{12} \tag{16}$$

However, because of the properties of the Fibonacci numbers $\beta_{11} = \beta_{22}$ and only the one new parameter and therefore one new frequency distribution need be calculated during each iteration.

Figure 39 illustrates the steps followed in this search procedure. Convergence is obtained when the difference between two consecutive values of β is less than some arbitrarily set number. Whichever of these β magnitudes is best according to the calibration procedure is taken as the parameter magnitude.

The adjustment process for A_i and B_j and the estimation of β_s are conducted in an inter-dependent way. The A_i and B_j are adjusted for each β_s before the calculation of the model estimated trip length

frequency distribution. The logic of this process is illustrated in Figure 40.

8.3/ Goodness of Fit Statistics

The calibration criterion used in the estimation of the gravity model parameters is the minimization of the differences between the observed and model estimated trip length frequency distributions. However, the trip interchange elements of the observed and model-estimated trip matrices may be compared in many ways and a number of alternative goodness of fit statistics were examined. These statistics included chi-squared, r-squared, standard deviation of the trip interchange residuals, likelihood ratios and a phi-statistic from information theory. A detailed appraisal of these alternative statistics is presented in Appendix C. It was concluded that the best goodness of fit measure to use is the phi-statistic. The main features of the appraisal presented in Appendix C are highlighted in the following paragraphs.

The behaviour of the alternative goodness of fit statistics was examined for eight of the larger Ontario census areas. This problem could not be approached in the traditional statistical sense since the magnitudes of the statistics are influenced by the number of zones, the total number of trip interchanges, and the characteristic probability distributions of the observed and estimated trip interchange magnitudes. A simulation approach was taken to evaluate the performance of alternative statistics.

Simulated trip interchange matrices were obtained by multiplying the observed trip interchange entries by a percentage error where these errors were

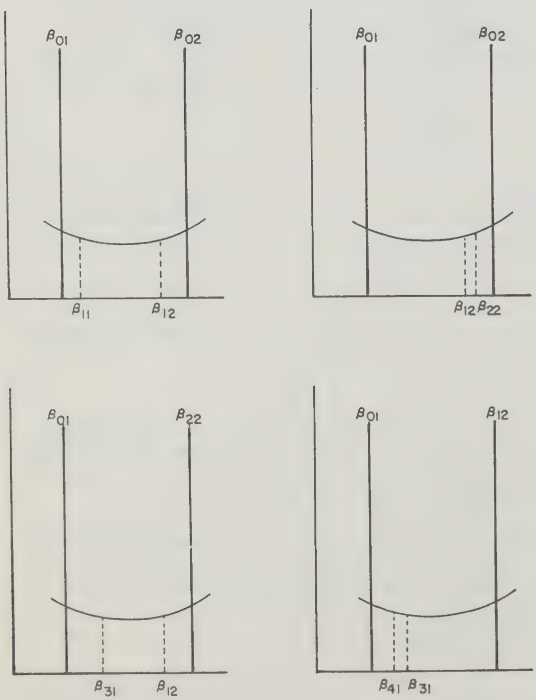


Figure 39/ Golden Section Search Procedure

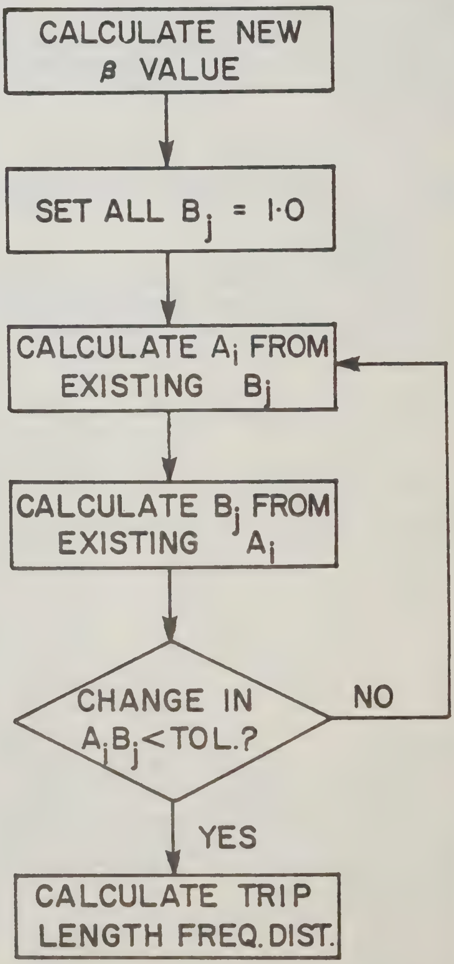


Figure 40/ Calibration Procedure for Doubly-Constrained Gravity Model

randomly generated from a rectangular distribution with mean of zero and a specified range. Six percentage ranges were used: 10, 25, 50, 75, 100 and 150. The simulated trip interchange matrices obtained in this way were then adjusted so that the row and column totals equalled the observed labour force and employment magnitudes. In this sense the simulated trip interchange magnitudes were similar to those produced by a doubly-constrained model.

The most commonly used goodness of fit measure in transport planning is the coefficient of determination, R^2, which is defined as:

$$R^2 = 1 - \frac{\sum_i \sum_j (T_{ij} - T^*_{ij})^2}{\sum_i \sum_j (T_{ij} - \bar{T}_{ij})^2} \qquad (17)$$

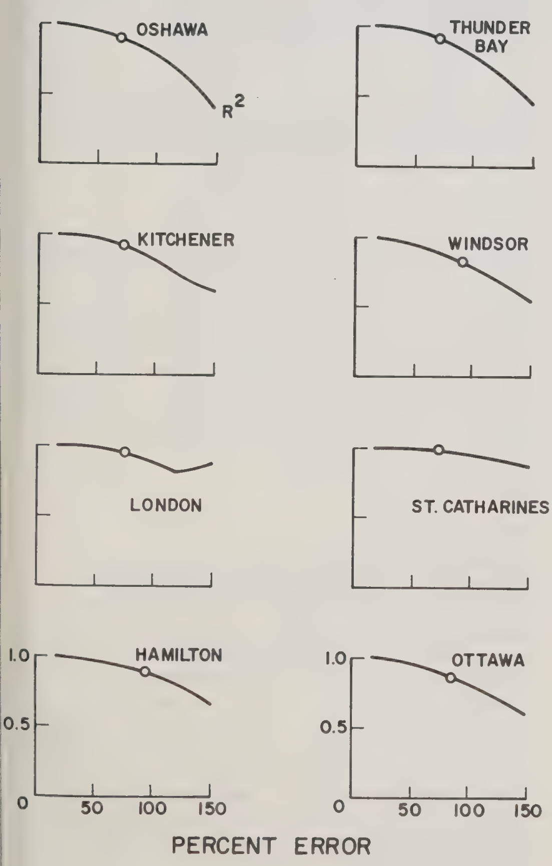


Figure 41/ Variation in R^2 with Allowable Percentage Error

where T_{ij} = observed number of trip interchanges between zones i and j.

Figure 41 shows the variation in R^2 magnitudes for the eight census areas with allowable percentage error. This diagram demonstrates clearly the insensitivity of R^2 to changes in the error magnitude. Even when the allowable percentage error is 100%, the R^2 magnitudes are still greater than about 0.7. The evidence suggests that there is little point in using R^2 to trace changes in model performance.

The phi-statistic of information theory is defined as:

$$\phi = \sum_i \sum_j T_{ij} \left| \log_e \frac{T^*_{ij}}{T_{ij}} \right| \qquad (18)$$

That is ϕ compares the ratio of the estimated and observed trip interchange magnitudes and weights this by the observed number of trips. Because of this, ϕ is most sensitive to errors in large interchanges. The larger the absolute value of phi, the poorer the degree of fit of the model.

Figure 42 shows the variation in the magnitude of the phi-statistics with increasing percentage error. It

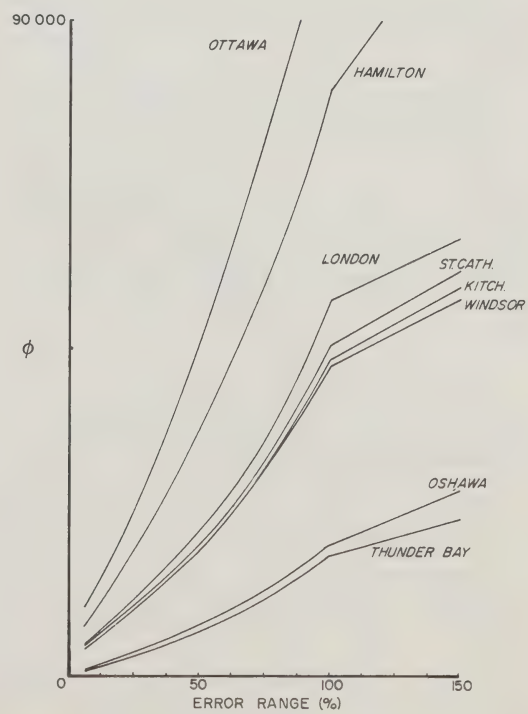


Figure 42/ Variation in Phi with Allowable Error Percentage

is important to note that up to error magnitudes of 100% the phi-statistic increases almost linearly with increasing percentage error. The phi-statistic is also consistent in its ranking of areas across all error magnitudes except for Kitchener and Windsor.

Figure 43 shows the phi-statistic divided by the number of interchange when the estimated matrix has been generated using a 75% error magnitude. This diagram shows that the transformed phi-statistic is quite consistent across all urban area sizes, which suggests that the transformed phi-statistic may be used as a general goodness of fit statistic.

8.4/ Production-Constrained Models

Table 23 summarizes the characteristics of the production-constrained models estimated for eight Ontario census areas. This table shows that the observed and simulated mean trip lengths are very close to each other despite the fact that the calibration criterion was minimization of the sum of the absolute differences between the ordinates of the observed and simulated trip length frequency distributions. The β magnitudes range from 0.152 in Windsor to 0.237 in the Kitchener census area and tend to decrease in magnitude with increasing mean trip length. The transformed phi magnitudes shown in the right-hand column of the table suggest that the goodness of fit of the production-constrained models is equivalent to a randomly introduced error of 75 to 100%.

8.5/Attraction-Constrained Models

The properties of the attraction-constrained gravity models for eight census areas are summarized in Table 24. The β magnitudes are all smaller than the β magnitudes of the production-constrained model. The β magnitudes tend to decrease with increasing mean trip length but at a faster rate than for the production-constrained models. The phi-statistic and the transformed phi-statistic indicate that that goodness of fit of the attraction-constrained models is not as high as for the production-constrained models.

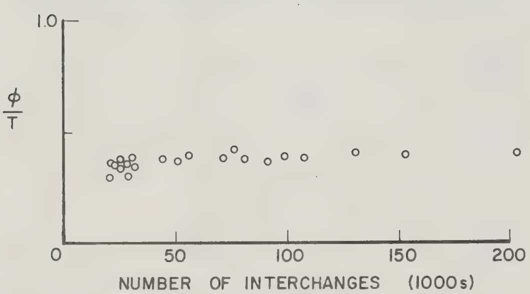


Figure 43/ Variation in ϕ/T with Number of Interchanges

Table 23/ Properties of Production-Constrained Gravity Models for Eight Census Areas

Census Area	β	Mean Trip Length (km) Obs.	Sim.	R^2	Phi	Phi/T
Oshawa	0.177	4.10	4.24	0.84	12 963	0.38
Thunder Bay	0.191	5.37	5.19	0.87	11 142	0.34
Kitchener	0.237	4.80	4.70	0.85	33 746	0.42
Windsor	0.152	7.59	7.39	0.75	36 900	0.49
St. Catharines	0.210	6.50	6.11	0.94	41 801	0.46
London	0.177	6.57	6.12	0.94	54 308	0.56
Hamilton	0.160	8.05	8.23	0.76	87 294	0.57
Ottawa	0.202	7.63	7.48	0.74	130 284	0.65

Table 24/ Properties of Attraction-Constrained Gravity Models for Eight Census Areas

Census Area	β	Mean Trip Length (km) Obs.	Sim.	R^2	Phi	Phi/T
Oshawa	0.138	4.10	4.22	0.83	13 720	0.40
Thunder Bay	0.172	5.37	5.13	0.84	12 563	0.38
Kitchener	0.230	4.80	4.69	0.86	34 439	0.42
Windsor	0.084	7.59	7.75	0.62	43 829	0.59
St. Catharines	0.201	6.50	6.06	0.94	44 385	0.48
London	0.146	6.57	6.60	0.92	53 358	0.55
Hamilton	0.102	8.05	0.57	0.58	103 136	0.68
Ottawa	0.122	7.63	7.57	0.68	151 152	0.75

8.6/ Doubly-Constrained Models

The doubly-constrained gravity model has been estimated for the 15 Ontario census areas and the properties of the calibrated models are summarized in Table 25. The observed and simulated mean trip lengths are close for most census areas. The phi and transformed phi-statistics shown in the table indicate that the doubly-constrained gravity model is superior to both the production-constrained and attraction-constrained versions of the model. The β parameter magnitudes are plotted against the observed mean trip length in Figure 44. While the β parameter decreases with increasing trip length, there is a significant amount of variation at any mean trip length. It should also be noted that the parameter magnitudes for the doubly-constrained model tend to be larger than the parameters for the singly-constrained models.

Table 25/ Properties of Doubly-Constrained Gravity Models for Fifteen Census Areas Using Distance Cost Functions

Census Area	β	Mean Trip Length (km) Obs.	Sim.	R^2	Phi	$\frac{Phi}{T}$
Guelph	0.165	3.76	3.80	0.86	5 742	0.29
Peterborough	0.288	3.30	3.11	0.92	5 709	0.28
Brantford	0.206	4.20	4.13	0.82	7 114	0.29
Sarnia	0.091	6.33	6.42	0.92	6 485	0.27
Kingston	0.105	5.68	5.86	0.90	8 281	0.31
Sault Ste. Marie	0.131	4.59	4.77	0.94	6 566	0.27
Oshawa	0.213	4.10	4.21	0.87	11 008	0.33
Thunder Bay	0.201	5.37	5.09	0.87	10 775	0.33
Sudbury	0.120	8.63	8.73	0.75	20 876	0.46
Kitchener	0.259	4.80	4.60	0.90	29 924	0.37
St. Catharines	0.214	6.50	6.05	0.96	37 359	0.41
Windsor	0.115	7.59	7.91	0.77	34 625	0.47
London	0.180	6.57	6.40	0.95	44 360	0.46
Hamilton	0.159	8.05	8.26	0.81	83 017	0.55
Ottawa	0.212	7.63	7.40	0.79	119 509	0.60

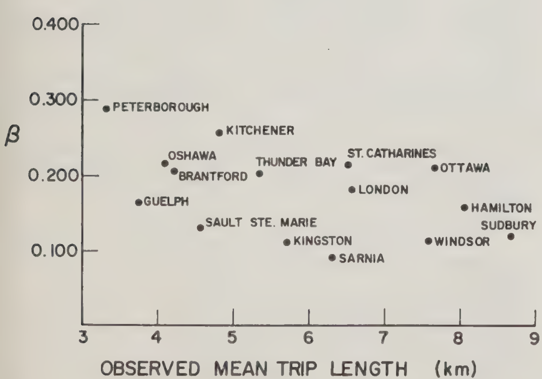


Figure 44/ Beta Parameter Magnitudes versus Observed Mean Trip Length

8.7/ Travel Time and Time-Distance Trip Cost Functions

It was mentioned earlier that the inter-zonal travel costs used in the gravity models calibrated in the previous sections were expressed in terms of road network distances in kilometres. Doubly-constrained models were calibrated for the Kitchener, Hamilton and Ottawa census area using travel times that were derived from the transport studies performed in those communities. The properties of the doubly-constrained gravity model estimated for those communities are summarized in Table 26.

Table 26/ Properties of Doubly-Constrained Gravity Models for Three Census Areas Using Travel Time Cost Functions

Census Area	β	Mean Trip Length (mins) Obs.	Sim.	R^2	Phi	$\frac{Phi}{T}$
Kitchener	0.150	10.23	9.79	0.74	36 212	0.45
Hamilton	0.121	13.38	13.48	0.76	88 430	0.58
Ottawa	0.119	13.51	13.71	0.78	121 352	0.60

This table indicates that the observed and simulated mean trip lengths are close. The largest differences are for the Kitchener census area. The phi-statistics indicate that the calibrated models are marginally inferior to those calibrated using network distances.

The doubly-constrained model was also estimated using travel costs that represented various combinations of inter-zonal distance and travel time for the Kitchener census area. The four generalized cost functions used are:

$$\text{I} \quad c^*_{ij} = \beta_{\text{distance}} \times c_{ij} + \beta_{\text{time}} \times tt_{ij} \qquad (19)$$

$$\text{II} \quad c^*_{ij} = \beta_{\text{distance}} \times c_{ij}\,\text{FACT} + \beta_{\text{time}} \times tt_{ij} \times (1\text{-FACT}) \qquad (20)$$

$$\text{III} \quad c^*_{ij} = c_{ij}/\text{Mean Trip Distance} + tt_{ij}/\text{Mean Trip Time} \qquad (21)$$

$$\text{IV} \quad c^*_{ij} = \left[c_{ij}/\text{Mean Trip Distance}\right] \times \text{FACT} + \left[tt_{ij}/\text{Mean Trip Time}\right](1\text{-FACT}) \qquad (22)$$

where

c^*_{ij} = "generalized" trip cost between zones i and j

β distance = magnitude estimated for the doubly-constrained model using network distances

β time 12 = magnitude estimated for the doubly-constrained model using network travel time

c_{ij} = network travel distance between zones i and j

tt_{ij} = network travel time between zones i and j

FACT is an index used to weight distance and speed differentially depending on the average travel speed between zones i and j

$$\text{FACT} = 1 \text{ if average speed} \geq 80 \text{ km/h}$$

$$= \frac{\text{average speed}-10}{70} \text{ if } 10 \leq \text{speed} \leq 80 \text{ km/h}$$

$$= 0 \text{ if average speed} \leq 10 \text{ km/h}$$

With the generalized cost function specified in Equation 19 the deterrence effects of distance and time are equally weighted. In Equation 20 when the average speed is greater than 80 km/h the deterrence effect is assumed to be all distance and when the average speed is less than 10 km/h the deterrence effect is assumed to be all travel time. In Equation 21 the distance and speed components are effectively normalized because they are divided by the area-wide mean trip distances and mean trip length. Equation 22 has the same effect as Equation 20.

Table 27 summarizes the properties of the doubly-constrained gravity models using the four different cost functions. These parameter magnitudes cannot be compared directly since the generalized travel cost units vary between each model. The observed and simulated mean trip lengths are close to each other but there is little change in the overall goodness of fit of the doubly-constrained model. The models are of better quality than those using travel times, but are marginally inferior to that estimated for the Kitchener census area using network travel distances as the travel costs.

8.8/ Comparisons of Residuals Across Model Types

It is useful to examine the trip interchange residuals in some detail in addition to the goodness of fit statistics. The trip interchange residuals for the Ottawa census area are examined for four separate model types in this section. The spatial distributions of residuals observed for Ottawa are broadly representative of the residuals for all census areas.

Figures 45 and 46 show the overestimation (simulated > observed) and underestimation (simulated < observed) residuals for the production-constrained gravity model estimated for Ottawa. Trips from the inner suburbs of Ottawa and the residential areas in Hull to the Ottawa central area employment zones tend to be overestimated by the production-constrained model. This tendency to overestimate trips to the central area is directly related to the heavy concentrations of employment in the Ottawa CBD. The underestimation residuals are primarily associated with trips to suburban employment locations in both Ottawa and Hull although there are some underestimation residuals associated with the longer trips to the CBD. This latter type of residual is probably associated with the timing of development of certain residential subdivisions and the growth of the federal government job opportunities in the central area.

Figures 47 and 48 illustrate the overestimation and underestimation residuals for the attraction-constrained gravity model estimated in Ottawa. The overestimation residuals tend to be clustered around the central employment area reflecting the tendency of the gravity model to allocate trips to the closest available opportunities.

The underestimation residuals are almost all associated with the long trips to the CBD employment opportunities. These residuals are a major cause of the poorer goodness of fit of the attraction-constrained model. The frequency distribution of the labour force by census tract is much more uniform than employment and inspection of the β parameter magnitudes in Tables 20 and 21 shows that the β parameter magnitude for the attraction constrained model falls in order to compensate.

Figures 49 and 50 show the overestimation and underestimation residuals for the doubly-constrained version of the gravity model. Figure 49 shows that the doubly-constrained model tends to reduce the number of overestimation residuals to the central area compared with the production-constrained model. On the other hand, Figure 50 shows that the underestimation residuals tend to increase for the doubly-constrained model compared with the production-constrained model.

Figures 51 and 52 show the overestimation and underestimation residuals for the doubly-constrained version of the gravity model using network travel times rather than network travel distances.

The comparison of the overestimation residuals plotted in Figure 51 with those plotted in Figure 49 illustrates the tendency of the model to further overestimate trips from some of the outer residential areas in Hull. Comparisons of the underestimation residuals in Figures 50 and 52 show that there is little change in the spatial pattern of these residuals.

The relative importance of the overestimated and underestimated residuals is further illustrated in Table 28 where the phi-statistics are summarized for the four gravity model types. This table clearly demonstrates that the major source of error is the underestimation of trip interchange magnitudes and that the superior behaviour of the doubly-constrained model is due to a reduction in the underestimation residuals. The table also demonstrates that the intrazonal underestimation residuals are a significant proportion of the total. However, in reviewing this decomposition of the residuals it should be recalled that the phi-statistic is more sensitive to underestimation errors than to overestimation errors.

Table 27/ Properties of the Doubly-Constrained Gravity Model for the Kitchener Census Area Using Different Cost Functions

Cost Function	β	Mean Trip Length Obs.	Mean Trip Length Sim.	R^2	Phi	$\frac{\text{Phi}}{\text{T}}$
I	0.490	2.77	2.72	0.85	31 123	0.38
II	0.998	1.50	1.46	0.82	32 366	0.40
III	0.300	4.80	4.56	0.85	31 563	0.39
IV	0.303	4.98	4.76	0.84	31 833	0.39

Table 28/ Relative Importance of Over- and Under-Estimation Residuals for Four Model Types for Ottawa

Gravity Model Type	Total Phi Over Est. Residuals	Total Phi Under Est. Residuals	Intra-Zonal Phi Over Est. Residuals	Intra-Zonal Phi Under Est. Residuals
Production - Constrained	23 634	106 650	426	21 597
Attraction - Constrained	21 717	129 435	0	30 055
Doubly - Constrained	21 049	98 460	873	16 622
Doubly - Constrained With Time	20 814	100 537	955	19 189

Table 29/ Goodness of Fit Statistics for Doubly-Constrained Gravity Models Using Different Travel Cost Functions for Kitchener Census Area

Travel Cost Function Type	Total Phi Over Est. Residuals	Total Phi Under Est. Residuals	Intra-Zonal Phi Over Est. Residuals	Intra-Zonal Phi Under Est. Residuals
Distance	8 178	22 875	485	3 008
Time	8 701	23 720	1 260	3 868
I	8 247	23 007	909	3 417
II	8 646	23 107	1 131	3 553
III	8 556	21 745	1 192	2 943
IV	8 752	27 510	1 173	3 231

8.9/ Comparison of Residuals for Alternative Cost Functions

A more detailed analysis of the doubly-constrained gravity models calibrated for the Kitchener census area using alternative travel cost functions is shown in Table 29. This table shows that there is little change between the different travel cost function types with the distance-time travel cost functions providing the best goodness of fit.

Figure 45/ Over-estimation Residuals for Production-Constrained Gravity Model in Ottawa

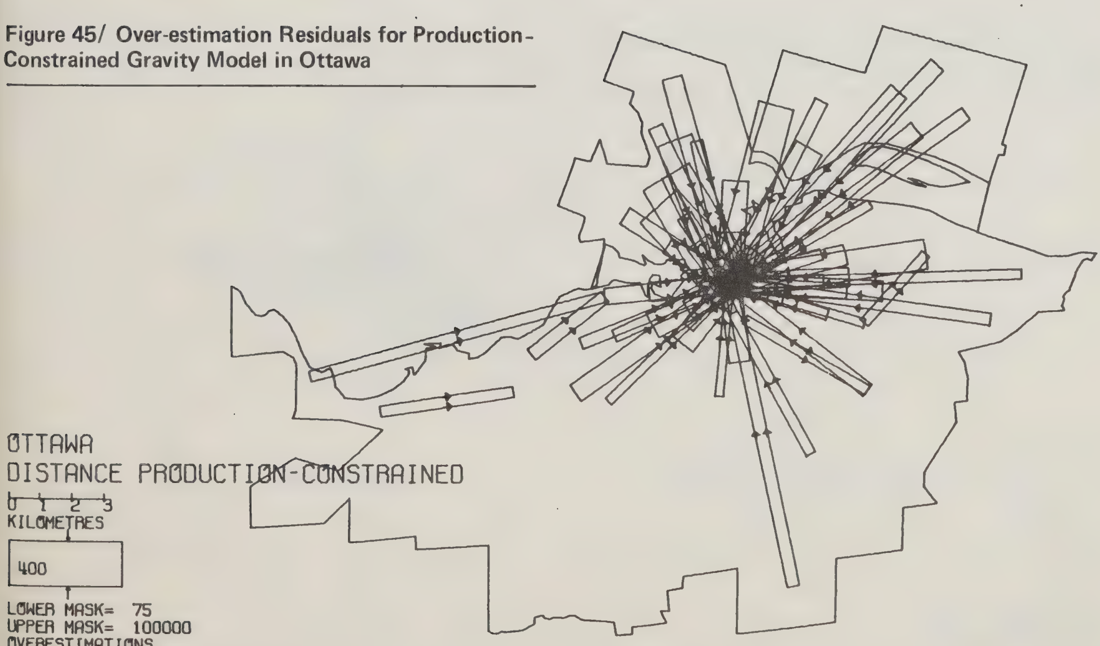


Figure 46/ Under-estimation Residuals for Production - Constrained Gravity Model in Ottawa

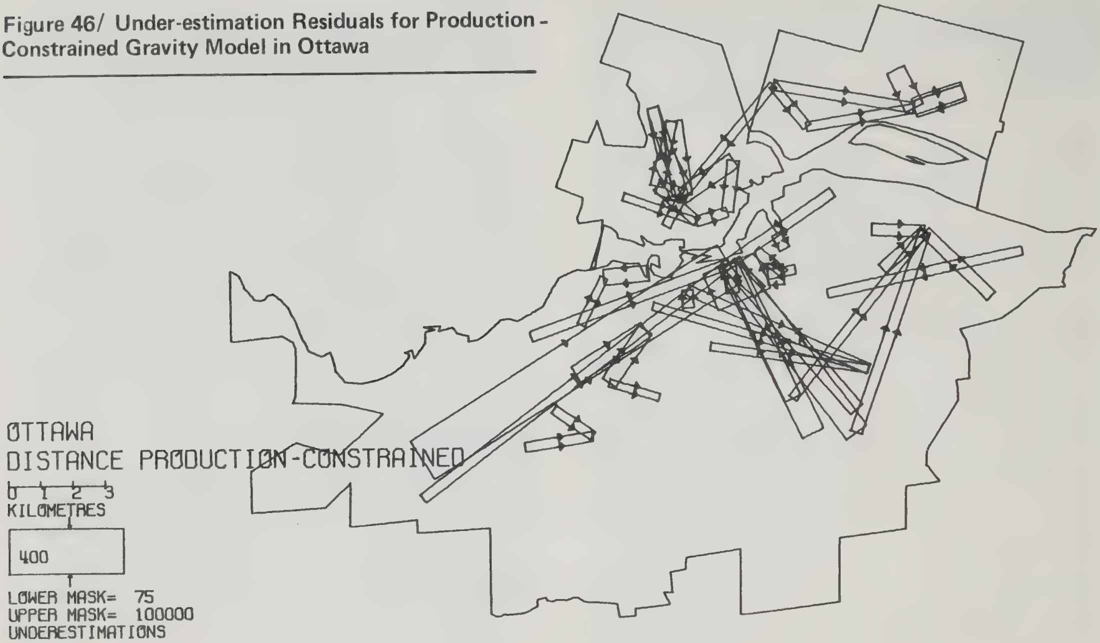


Figure 47/ Over-estimation Residuals for Attraction - Constrained Gravity Model in Ottawa

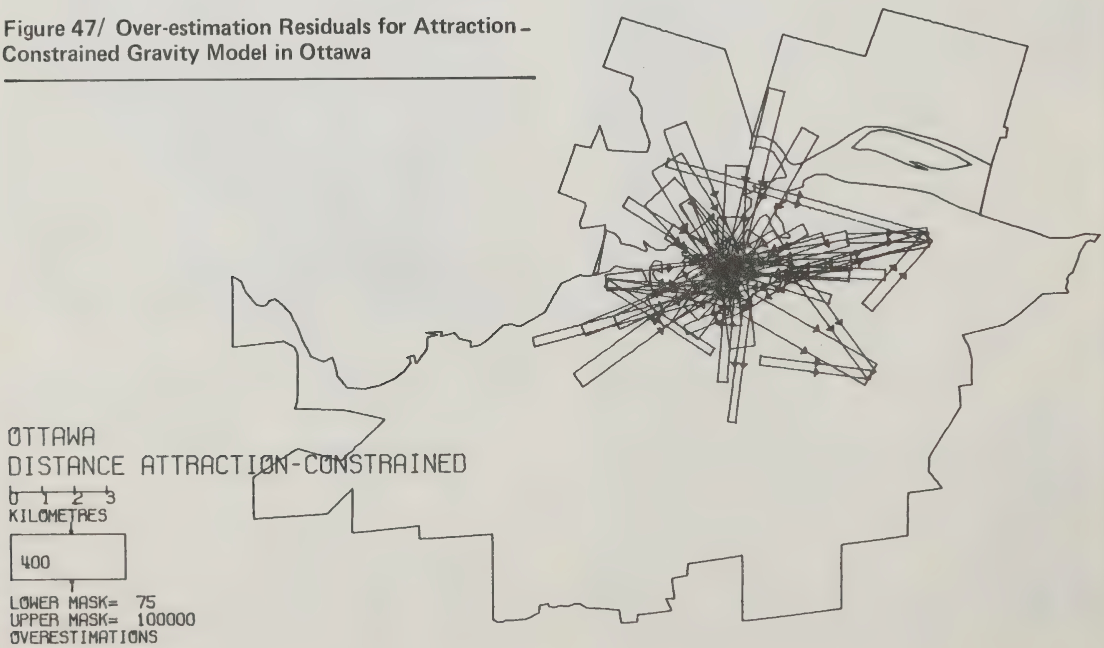


Figure 48/ Under-estimation Residuals for Attraction-Constrained Gravity Model in Ottawa

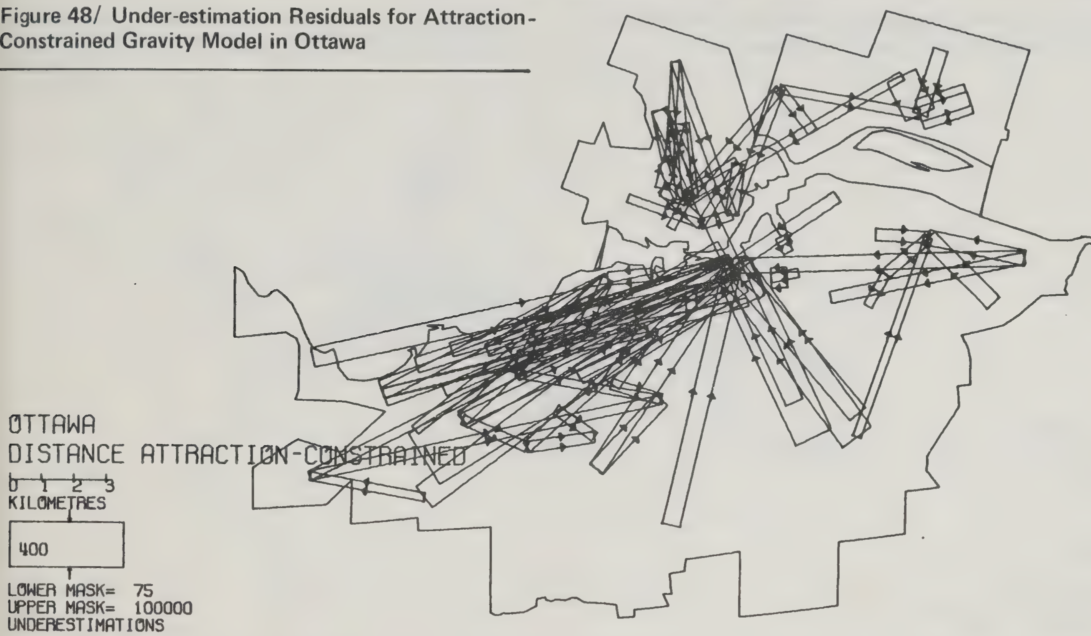


Figure 49/ Over-estimation Residuals for Doubly-Constrained Gravity Model in Ottawa

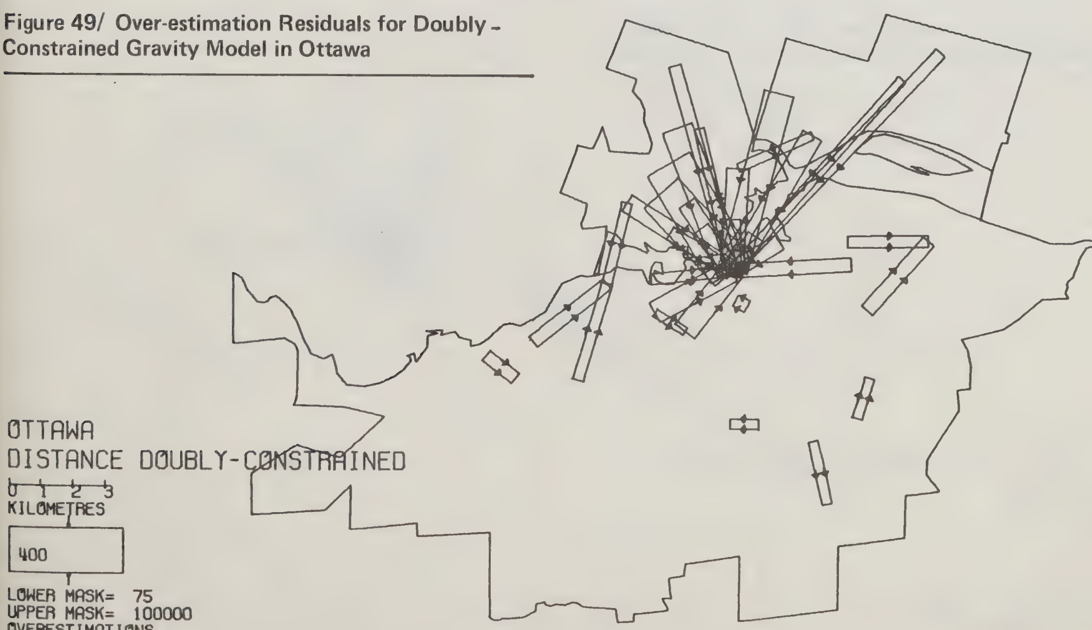


Figure 50/ Under-estimation Residuals for Doubly-Constrained Gravity Model in Ottawa

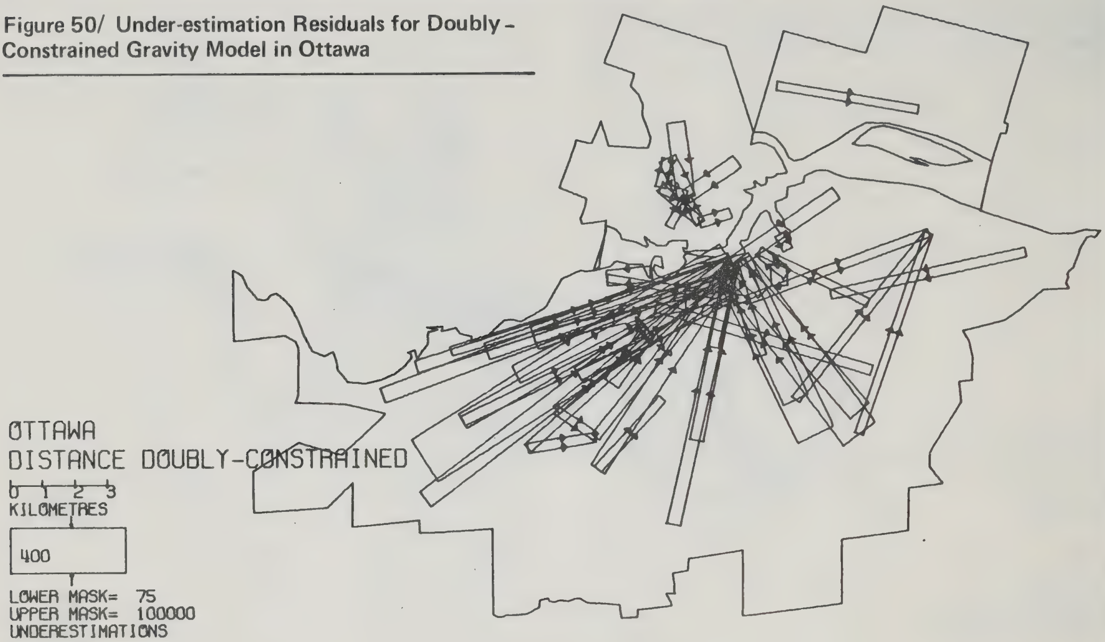


Figure 51/ Over-estimation Residuals for Doubly-Constrained Gravity Model in Ottawa Using Travel Times

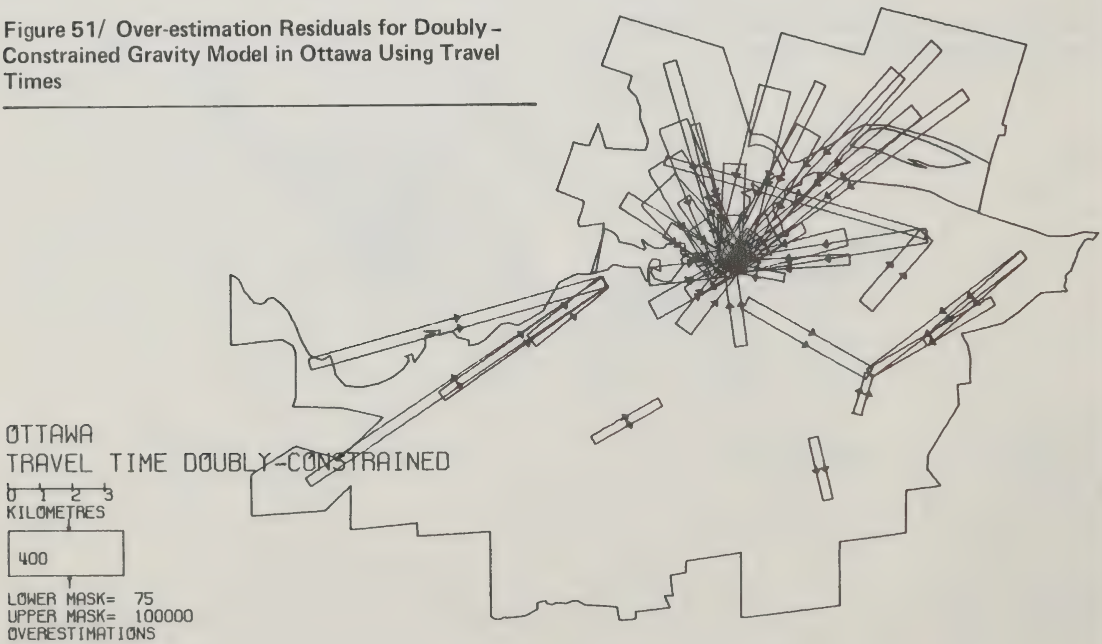
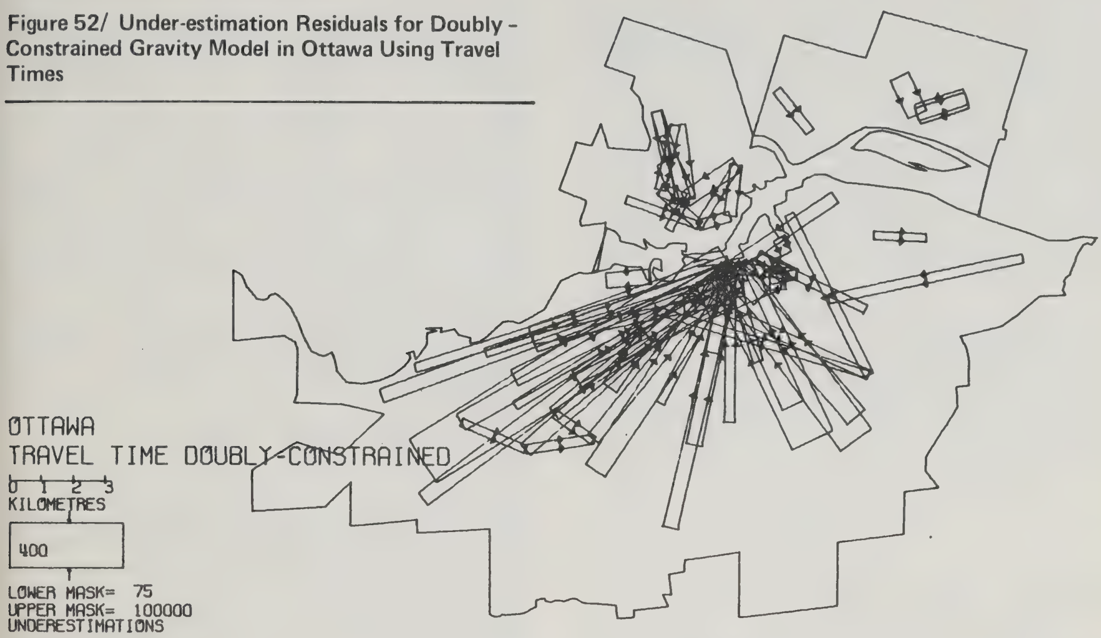


Figure 52/ Under-estimation Residuals for Doubly - Constrained Gravity Model in Ottawa Using Travel Times



9/ Stratified Trip Distribution Models

This section examines the qualities of a number of stratified gravity models. The first group of stratified models examined are multi-parameter gravity models in which separate parameter magnitudes are estimated for each origin zone in production-constrained, attraction-constrained and doubly-constrained forms of the gravity mode. The second group of stratified gravity models discussed have separate parameter magnitudes estimated for from four to eight calibration subregions where the census tracts in each of these calibration subregions have been identified by the clustering procedure described in Section 6. The final set of gravity models evaluated in this section are stratified by socio-economic group.

9.1/ Multi-Parameter Production-Constrained Models

Table 30 summarizes the characteristics of the multi-parameter production-constrained gravity models estimated for eight census areas. With these models separate parameter magnitudes were estimated for each residential zone. The table shows that in all cases the area-wide observed and simulated mean trip lengths are close to each other for all census areas. The multi-community census areas of Kitchener and St. Catharines exhibit the largest differences between the observed and simulated trip lengths. A comparison of the goodness of fit statistics in Table 30 with those presented in Table 23 for the single parameter production-constrained models shows that the multi-parameter models all exhibit superior goodness of fit characteristics. The largest improvements were in the Kitchener and St. Catharines census areas.

Figure 53 shows the spatial distribution of origin-zone specific parameters for the Kitchener census area. It should be recalled that a β parameter magnitude of 0.237 was obtained for the aggregate production-constrained model; the range obtained for the multi-parameter model was from 0.077 to 0.479. Figure 53 illustrates that most of the fringe residential zones not located close to employment areas have lower than average β magnitudes, while those for the inner residential areas closer to the large employment zones have much larger values. The calibrated magnitudes for each residential zone reflect not only the mean trip length characteristics of the zone but also the proximity of the zone to large concentrations of trip attractions.

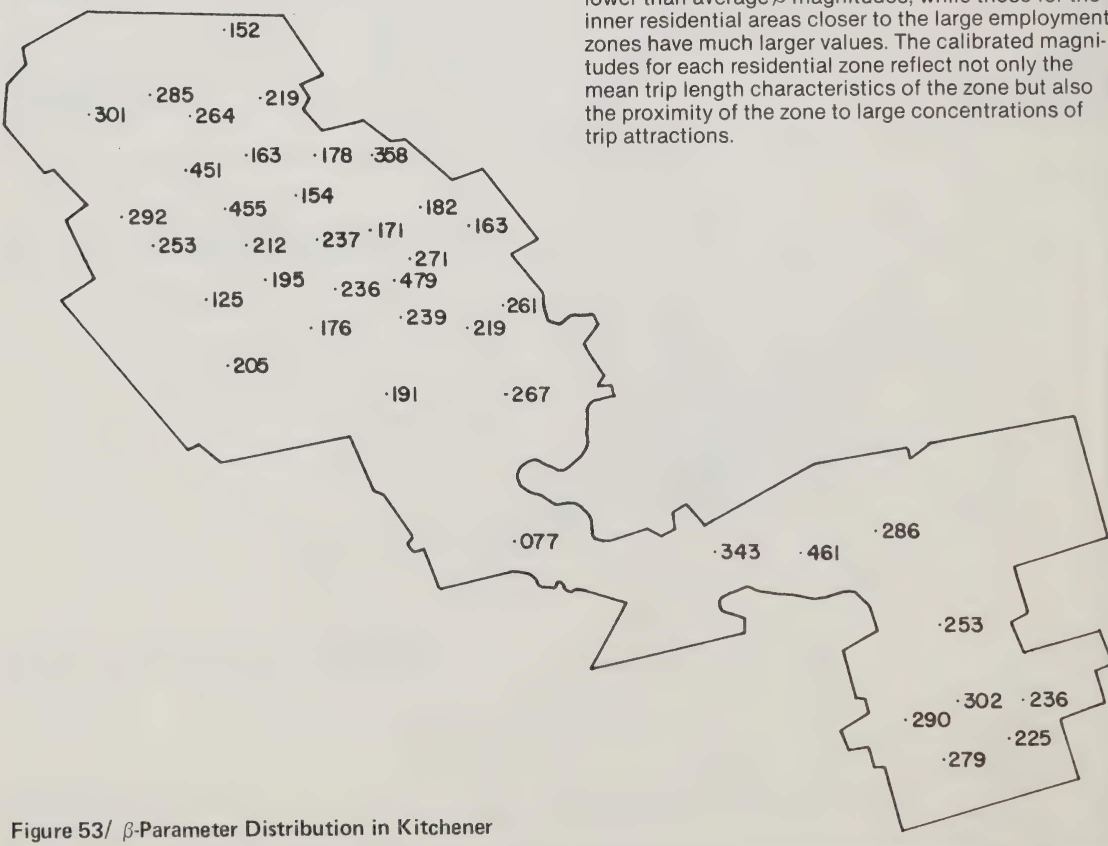


Figure 53/ β-Parameter Distribution in Kitchener

Table 30/ Properties of the Multi-Parameter Production-Constrained Gravity Models for Eight Census Areas

Census Area	β Range	Mean Trip Length (km) Obs.	Sim.	R^2	Phi	$\frac{Phi}{T}$
Oshawa	.003 - .472	4.10	4.16	0.90	10 827	0.32
Thunder Bay	.032 - .314	5.37	5.21	0.88	10 775	0.33
Kitchener	.077 - .479	4.80	4.59	0.89	30 420	0.37
Windsor	.015 - .414	7.59	7.52	0.83	33 903	0.46
St. Catharines	.127 - .367	6.50	5.97	0.97	36 164	0.39
London	.030 - .497	6.57	6.42	0.95	49 043	0.50
Hamilton	.032 - .489	8.05	7.92	0.87	80 669	0.53
Ottawa	.047 - .497	7.63	7.33	0.75	126 048	0.63

Figure 54 shows the spatial distribution of origin-zone specific β parameter magnitudes for the London census area. The β magnitude for the aggregate production-constrained model was 0.177 and the range illustrated in Figure 54 is from 0.030 to 0.497. Again, the lowest magnitudes are on the urban fringe and the largest values in zones close to the CBD and the University of Western Ontario.

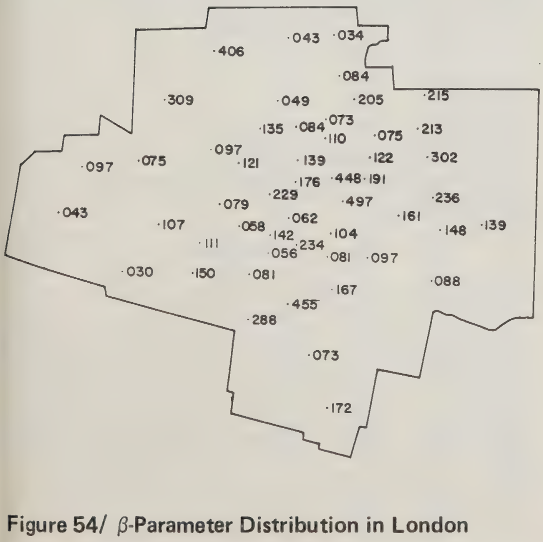


Figure 54/ β-Parameter Distribution in London

9.2/ Multi-Parameter Attraction-Constrained Models

Table 31 summarizes the properties of the multi-parameter attraction-constrained models estimated for eight Ontario census areas. A comparison of the goodness of fit statistics in Table 31 with those in Table 24 indicates that the multi-parameter attraction-constrained model has poorer goodness of fit characteristics than the aggregate model. The only census areas in which improvements occur are Thunder Bay, Kitchener and St. Catharines, the census areas with strong multi-community spatial structures.

9.3/ Multi-Parameter Doubly-Constrained Models

Table 32 summarizes the properties of the multi-parameter doubly-constrained models estimated for the eight Ontario census areas. A comparison of the goodness of fit statistics of Table 32 with those of Table 25 indicates that the multi-parameter doubly-constrained model is superior to the single parameter doubly-constrained model. In addition, the doubly-constrained multi-parameter model is superior to the other two multi-parameter models.

9.4/ Subregion Specific Models

Table 33 summarizes the properties of production-constrained gravity models estimated with separate β parameters for a number of individual calibration sub-regions for eight census areas. The cluster analyses described in Section 7 formed the basis for the calibration subregions isolated for each census area. This table shows that the observed and simulated mean trip lengths are close together for all census areas. A comparison of the goodness of fit statistics in Table 34 with those in Tables 23, 25 and 30 shows that the subregion specific model is superior to the single parameter production-constrained model and inferior to both the single parameter doubly-constrained and multi-parameter production-constrained models.

Table 34 summarizes the results obtained from the calibration of doubly-constrained models incorporating specific β parameters for the same calibration subregions used for the production-constrained models summarized in Table 33. A comparison of the goodness of fit statistics of Table 34 with those in Tables 25 and 33 shows that the doubly-constrained subregion specific models performed at about the same level or marginally better than the single parameter doubly-constrained models but were inferior to the multi-parameter doubly-constrained models.

9.5/ Models Stratified by Socio-Economic Group

A number of gravity models were estimated using the stratified journey-to-work tabulations obtained from Statistics Canada (Section 3). Table 35 summarizes the characteristics of production-constrained models estimated for members of the labour force without a car available and with a car available for the journey-to-work. The only goodness of fit statistics reported in

Table 35 are the coefficients of determination since these production-constrained models were calibrated in the first research project. Inspection of Table 35 shows that there is generally good agreement between the observed and simulated mean trip lengths for the two groups and that in most census areas the β parameter magnitude for the non-car owners is larger than for car owners.

Table 36 summarizes the characteristics of doubly-constrained models estimated for non car owners in three census areas using both network distances and network times. Table 37 summarizes the information for the models calibrated for car owners. Since the zone systems and numbers of linkages are fewer for the stratified models it is appropriate to compare the transformed phi-statistics in Tables 36 and 37 with those in Table 25. This comparison suggests that the doubly-constrained models stratified by car ownership status perform at a superior level to the doubly-constrained single parameter models. The β parameter magnitudes for the captive members of the labour force are significantly higher than for the noncaptives for all three census areas, thus reflecting the greater sensitivity of captives to travel costs.

Tables 38, 39 and 40 summarize the characteristics of the stratified models calibrated for home owners and home renters. Tables 41, 42 and 43 summarize the characteristics of the stratified models calibrated for periods of residence of greater than six years and for five years and less. The comparison of home owner and home renter behaviour indicates that the differences are not systematic in terms of both the mean trip length comparisons and the β parameter comparisons. On the other hand, the period of residence comparisons show that there are significant differences between the two groups. The β parameters in the three census areas for periods of residence of six years and greater are larger than for the shorter periods of residence. The transformed phi-statistics show that this stratification has greater explanatory power than the single parameter model and is marginally inferior to the stratification based on car ownership status.

9.6/ Comparison of Model Stratification

The two stratification methods used in this section were the origin end stratification of the deterrence function parameter, which allows for a real variation in the mean trip length, and the socio-economic stratification of the trip linkage matrix, which produces a separate model for each group. The use of a stratified deterrence function generally produced a marginal improvement in model performance and a greater improvement was observed when models were calibrated for separate socio-economic groups. This improvement is due both to the effects of a more representative deterrence function for the separate groups, and to the additional information introduced into the model in the form of stratified trip ends.

Table 31/ Properties of Multi-Parameter Attraction-Constrained Gravity Models for Eight Census Areas

Census Area	β Range	Mean Trip Length (km) Obs.	Sim.	R^2	Phi	$\frac{\text{Phi}}{\text{T}}$
Oshawa	.073 - .414	4.10	4.04	0.73	18 004	0.53
Thunder Bay	.118 - .497	5.37	5.21	0.84	11 717	0.35
Kitchener	.127 - .492	4.80	4.61	0.85	33 562	0.41
Windsor	.045 - .497	7.59	7.29	0.64	44 577	0.60
St. Catharines	.103 - .409	6.50	5.69	0.96	41 108	0.45
London	.073 - .497	6.57	6.49	0.89	56 708	0.58
Hamilton	.073 - .497	8.05	7.58	0.58	112 857	0.74
Ottawa	.073 - .497	7.63	6.87	0.68	144 116	0.72

Table 32/ Properties of Multi-Parameter Doubly-Constrained Gravity Models for Eight Census Areas

Census Area	β Range	Mean Trip Length (km) Obs.	Sim.	R^2	Phi	$\frac{\text{Phi}}{\text{T}}$
Oshawa	.056 - .461	4.10	4.16	0.91	10 204	0.30
Thunder Bay	.094 - .354	5.37	5.28	0.88	9 922	0.30
Kitchener	.118 - .438	4.80	4.54	0.92	27 918	0.34
Windsor	.034 - .480	7.59	7.63	0.84	32 521	0.44
St. Catharines	.118 - .420	6.50	6.09	0.97	34 102	0.37
London	.072 - .491	6.57	6.36	0.96	43 133	0.44
Hamilton	.045 - .483	8.05	7.97	0.88	79 285	0.52
Ottawa	.099 - .493	7.63	7.38	0.83	113 143	0.56

Table 33/ Properties of Sub-Region Specific Parameter Production Constrained Gravity Models of Eight Census Areas

Census Area	No.	β Range	Mean Trip Length (km) Obs.	Sim.	R^2	Phi	$\frac{\text{Phi}}{\text{T}}$
Oshawa	4	.046 - .369	4.10	4.29	0.88	11 434	0.34
Thunder Bay	3	.129 - .206	5.37	5.20	0.87	11 293	0.34
Kitchener	3	.182 - .313	4.80	4.55	0.87	33 023	0.41
Windsor	5	.098 - .414	7.59	7.44	0.75	36 589	0.49
St. Catharines	3	.188 - .223	6.50	5.95	0.95	41 223	0.45
London	4	.073 - .204	6.57	6.61	0.94	51 310	0.53
Hamilton	7	.058 - .229	8.05	8.03	0.82	86 310	0.57
Ottawa	8	.139 - .407	7.63	7.44	0.74	125 013	0.62

Table 34/ Properties of Sub-Region Specific Parameter Doubly-Constrained Gravity Models of Eight Census Areas

Census Area	No.	β Range	Mean Trip Length (km) Obs.	Sim.	R^2	Phi	Phi/T
Oshawa	4	.152 - .377	4.10	4.03	0.88	11 051	0.33
Thunder Bay	3	.073 - .223	5.37	5.44	0.87	10 127	0.31
Kitchener	3	.249 - .292	4.80	4.50	0.90	30 368	0.37
Windsor	5	.115 - .497	7.59	7.70	0.77	34 594	0.46
St. Catharines	3	.191 - .309	6.50	5.85	0.96	38 384	0.42
London	4	.155 - .208	6.57	6.48	0.94	44 273	0.46
Hamilton	7	.118 - .191	8.05	8.16	0.83	83 120	0.54
Ottawa	8	.116 - .309	7.63	7.60	0.80	116 139	0.58

Table 35/ Production-Constrained Gravity Models for No Car Available and Car Available Groups for Fifteen Census Areas

Census Area	Group	β	Mean Trip Length (km) Obs	Sim	R^2
Guelph	No Car	0.10	3.40	3.60	0.73
	Car	0.14	3.90	3.90	0.78
Peterborough	No Car	0.33	3.00	2.80	0.80
	Car	0.22	3.40	2.24	0.88
Sarnia	No Car	0.08	4.50	5.10	0.68
	Car	0.05	6.70	6.98	0.87
Brantford	No Car	0.32	3.50	3.24	0.80
	Car	0.25	4.40	4.08	0.77
Sault Ste. Marie	No Car	0.23	3.50	3.62	0.79
	Car	0.20	4.90	4.98	0.92
Kingston	No Car	0.07	4.30	4.56	0.83
	Car	0.10	6.00	6.10	0.85
Thunder Bay	No Car	0.26	4.60	4.30	0.89
	Car	0.18	5.60	5.40	0.92
Oshawa	No Car	0.15	3.90	4.10	0.74
	Car	0.22	4.10	4.10	0.80
Sudbury	No Car	0.22	6.40	5.78	0.72
	Car	0.10	9.20	9.20	0.68
Kitchener	No Car	0.32	4.10	3.84	0.83
	Car	0.22	5.10	5.04	0.83
Windsor	No Car	0.17	6.00	6.20	0.76
	Car	0.15	8.00	7.83	0.82
London	No Car	0.22	5.10	5.20	0.90
	Car	0.21	6.90	6.10	0.88
St. Catharines	No Car	0.36	4.70	4.14	0.96
	Car	0.23	6.80	5.90	0.94
Hamilton	No Car	0.25	5.80	5.70	0.82
	Car	0.16	8.80	8.70	0.91
Ottawa	No Car	0.23	6.10	6.30	0.79
	Car	0.17	8.30	8.30	0.75

Table 36/ Properties of Doubly-Constrained Gravity Models Calibrated Using Travel Time and Travel Distance For Three Urban Areas: No Car Available

Census Area and Deterrence Measure	β	Mean Trip Length (km, min) Obs.	Sim.	R^2	Phi	Phi/T
Kitchener						
Distance	0.386	4.10	3.63	0.88	10 235	0.36
Time	0.181	9.02	8.87	0.82	11 871	0.42
Hamilton						
Distance	0.237	5.83	5.80	0.87	21 563	0.42
Time	0.169	10.41	10.54	0.81	23 577	0.46
Ottawa						
Distance	0.248	6.15	6.18	0.84	35 789	0.49
Time	0.156	11.66	11.84	0.85	35 355	0.48

Table 37/ Properties of Doubly-Constrained Gravity Models Calibrated Using Travel Time and Travel Distance For Three Urban Areas: Car Available

Census Area and Deterrence Measure	β	Mean Trip Length (km, min) Obs.	Sim.	R^2	Phi	Phi/T
Kitchener						
Distance	0.239	5.07	4.91	0.90	16 310	0.29
Time	0.141	10.61	10.48	0.83	19 211	0.34
Hamilton						
Distance	0.147	8.79	9.01	0.97	35 102	0.33
Time	0.113	13.70	14.13	0.84	42 207	0.39
Ottawa						
Distance	0.158	8.34	8.51	0.80	60 604	0.46
Time	0.113	14.31	14.44	0.82	57 245	0.44

Table 38/ Production-Constrained Gravity Models for Home Owner and Home Renter Groups for Fifteen Census Areas

Census Area	Group	β	Mean Trip Length (km) Obs	Sim	R^2
Guelph	Owners	0.14	3.60	3.80	0.77
	Renters	0.11	3.80	3.90	0.75
Peterborough	Owners	0.29	3.30	3.10	0.86
	Renters	0.26	3.00	2.80	0.85
Sarnia	Owners	0.07	6.40	6.80	0.82
	Renters	0.02	5.70	5.90	0.85
Brantford	Owners	0.29	4.20	3.90	0.82
	Renters	0.21	3.80	3.60	0.75
Sault Ste. Marie	Owners	0.16	4.60	4.60	0.89
	Renters	0.08	4.10	4.30	0.83
Kingston	Owners	0.13	5.80	5.60	0.92
	Renters	0.08	5.20	5.20	0.81
Thunder Bay	Owners	0.20	5.40	5.20	0.91
	Renters	0.22	4.60	4.40	0.88
Oshawa	Owners	0.23	4.10	4.10	0.78
	Renters	0.16	3.70	3.90	0.80
Sudbury	Owners	0.13	8.60	8.40	0.71
	Renters	0.20	7.90	5.90	0.71
Kitchener	Owners	0.28	4.70	4.40	0.83
	Renters	0.26	4.70	4.30	0.84
Windsor	Owners	0.15	7.70	7.60	0.81
	Renters	0.16	6.70	6.30	0.80
London	Owners	0.21	6.60	6.20	0.88
	Renters	0.19	5.90	5.30	0.91
St. Catharines	Owners	0.28	6.30	5.30	0.95
	Renters	0.24	6.10	5.30	0.93
Hamilton	Owners	0.19	8.30	8.00	0.90
	Renters	0.19	6.90	6.70	0.84
Ottawa	Owners	0.20	8.40	8.20	0.77
	Renters	0.15	6.40	6.80	0.77

Table 39/ Properties of Doubly-Constrained Gravity Models Calibrated Using Travel Time and Travel Distance For Three Urban Areas: Home Owners

Census Area and Deterrence Measure	β	Mean Trip Length (km, min) Obs.	Sim.	R^2	Phi	$\frac{\text{Phi}}{\text{T}}$
Kitchener						
Distance	0.310	4.77	4.30	0.89	18 131	0.34
Time	0.162	10.15	9.80	0.81	20 116	0.37
Hamilton						
Distance	0.170	8.26	8.36	0.91	37 954	0.35
Time	0.126	13.13	13.52	0.83	44 016	0.40
Ottawa						
Distance	0.195	8.43	8.29	0.80	55 649	0.49
Time	0.110	14.22	14.95	0.82	50 916	0.45

Table 40/ Properties of Doubly-Constrained Gravity Models Calibrated Using Travel Time and Travel Distance For Three Urban Areas: Home Renters

Census Area and Deterrence Measure	β	Mean Trip Length (km, min) Obs.	Sim.	R^2	Phi	$\frac{\text{Phi}}{\text{T}}$
Kitchener						
Distance	0.306	4.67	4.14	0.88	10 600	0.35
Time	0.144	9.90	9.97	0.83	11 390	0.38
Hamilton						
Distance	0.191	6.88	6.67	0.86	21 039	0.42
Time	0.123	11.56	12.10	0.80	22 565	0.45
Ottawa						
Distance	0.196	6.43	6.59	0.82	42 722	0.48
Time	0.139	12.24	12.13	0.84	41 002	0.46

Table 41/ Production-Constrained Gravity Models for 6+ Years Resident and 0-5 Years Resident Groups for Fifteen Urban Areas

Census Area	Group	β	Mean Trip Length (km) Obs	Sim	R^2
Guelph	6+ Yrs.	0.10	3.50	3.70	0.74
	0-5	0.13	3.90	3.90	0.78
Peterborough	6+	0.21	3.20	3.20	0.87
	0-5	0.18	3.30	3.30	0.84
Sarnia	6+	0.12	6.00	5.80	0.85
	0-5	0.03	6.40	6.70	0.88
Brantford	6+	0.39	4.10	3.40	0.75
	0-5	0.22	4.20	3.90	0.79
Sault Ste. Marie	6+	0.16	4.50	4.50	0.89
	0-5	0.12	4.50	4.60	0.84
Kingston	6+	0.19	5.40	5.00	0.89
	0-5	0.07	5.60	5.90	0.79
Thunder Bay	6+	0.19	5.30	5.20	0.90
	0-5	0.21	5.20	4.90	0.91
Oshawa	6+	0.23	4.10	4.10	0.79
	0-5	0.16	3.90	3.90	0.80
Sudbury	6+	0.19	8.20	6.70	0.66
	0-5	0.11	8.70	8.50	0.67
Kitchener	6+	0.32	4.50	4.00	0.81
	0-5	0.24	4.80	4.70	0.87
Windsor	6+	0.18	7.10	6.80	0.80
	0-5	0.13	7.90	7.90	0.78
London	6+	0.23	6.10	5.90	0.87
	0-5	0.19	6.60	5.80	0.90
St. Catharines	6+	0.29	6.00	5.10	0.96
	0-5	0.22	6.10	6.00	0.93
Hamilton	6+	0.21	7.90	7.40	0.88
	0-5	0.19	7.90	7.60	0.87
Ottawa	6+	0.23	7.10	6.90	0.76
	0-5	0.16	7.90	8.00	0.78

Table 42/ Properties of Doubly-Constrained Gravity Models Calibrated Using Travel Time and Travel Distance for Three Urban Areas: 6+ Year Residents

Census Area and Deterrence Measure	β	Mean Trip Length (km , min) Obs	Sim	R^2	Phi	$\frac{Phi}{T}$
Kitchener						
Distance	0.337	4.57	4.07	0.88	14 059	0.37
Time	0.177	9.87	9.32	0.81	15 150	0.40
Hamilton						
Distance	0.207	7.74	7.44	0.90	31 012	0.39
Time	0.128	12.68	13.24	0.81	34 742	0.43
Ottawa						
Distance	0.223	7.10	7.06	0.81	42 337	0.48
Time	0.135	12.72	13.04	0.22	40 507	0.46

Table 43/ Properties of Doubly-Constrained Gravity Models Calibrated Using Travel Time and Travel Distance for Three Urban Areas: 0-5 Year Residents

Census Area and Deterrence Measure	β	Mean Trip Length (km , min) Obs	Sim	R^2	Phi	$\frac{Phi}{T}$
Kitchener						
Distance	0.249	4.83	4.68	0.91	13 187	0.29
Time	0.144	10.18	10.12	0.84	15 837	0.34
Hamilton						
Distance	0.151	7.92	8.13	0.88	29 965	0.38
Time	0.116	12.59	13.08	0.82	32 210	0.41
Ottawa						
Distance	0.164	7.91	8.01	0.81	54 848	0.47
Time	0.119	13.84	13.85	0.82	51 278	0.44

10/ Conclusions and Recommendations

Comparative analyses of home to work travel in the Ontario census areas have shown that the per capita journey-to-work distance varies from 1.24 km in the Peterborough CA to 3.93 km in the Toronto CMA. These per capita variations in the home to work travel distance are also reflected in variations in the mean home to work distance, which varies from 3.6 km in Peterborough to 10.8 km in the Toronto CMA.

While the per capita journey-to-work amount of travel tends to increase with increasing census area population there are significant deviations from this trend. At the 80 000 population scale, for example, the per capita amount of travel in Sarnia is about 32% higher than the mean for the six census areas of this size. These differences in the amount of travel are due to variations in spatial structure. Spatial structure is influenced strongly by the employment base of a community, topographic constraints and the location of desirable residential areas. Significant opportunities exist for reducing work trip travel through enlightened spatial management.

The spatial variations in the amount of residential labour force may be predicted with very high reliability from a knowledge of the dwelling unit composition of each census tract. The dwelling unit composition of each census tract may also be used to predict the number of transit captives and transit non-captives in each census tract. It is recommended that work trip production studies in Ontario municipalities be standardized in terms of dwelling unit composition. Equivalent trip attraction rates cannot be established from the census data since independent employment and land use measures are not available.

As an extension of these Ontario-wide trip generation studies it is recommended that centralized assessment records be analyzed to establish Ontario-wide dwelling unit and employment-end land use classifications. Centralized assessment records may also be used to update dwelling unit labour force composition rates.

Bi-proportional matrix balancing techniques and cluster analysis methods may be used to establish the commuting structure of a community. The bi-proportional techniques allow the pure interaction effects to be isolated that are independent of zone trip end magnitudes. Cluster analysis has been used to establish the commutersheds on a consistent basis for all Ontario census areas. These commutershed analyses demonstrated that there are five broad determinants of commuting in the Ontario census areas and these are:

(i) municipal boundaries;
(ii) topographic and man-made constraints;
(iii) the time sequence of urban development;
(iv) socio-economic factors; and
(v) the domination of large employment centres.

The most common and important factors are the timing of development and socio-economic factors.

Current goodness of fit measures such as the coefficient of determination and chi-square are unsuitable for assessing the quality of calibrated trip distribution models. It is recommended that the phi-statistic of information theory be used to measure the goodness of fit of trip distribution models. The phi-statistic divided by the number of linkages in an area provides a statistic that may be used to compare model goodnesses of fit across urban areas of different sizes.

Production-constrained, attraction-constrained and doubly-constrained gravity models have been calibrated for selected census areas and these studies have shown that the doubly-constrained model is superior to the two singly-constrained models. However, the phi-statistics showed that the doubly-constrained models had goodnesses of fit equivalent to a trip table produced by a random error of 75 to 100%. The travel deterrence parameter tended to decrease with increasing population size but there is a great deal of variation to this trend. However, the relationship may be used to establish initial estimates of the most appropriate parameter magnitudes.

Doubly-constrained gravity models using network travel times rather than network distances estimated for the Kitchener, Hamilton and Ottawa census areas were marginally inferior to those using network distances. Inspection of the origin-zone specific goodness of fit statistics for the Ottawa CMA showed that the network time based model improved the goodness of fit of some of the outlying residential census tracts.

Stratified trip distribution models have been calibrated in which separate travel deterrence parameters were estimated for selected geographic subregions of the census areas and different socio-economic groups. Multi-parameter production-constrained models were all superior to the single parameter-production constrained models but were not as good as the single parameter doubly constrained models.

The multi-parameter doubly-constrained models were all superior to the single parameter models. The most useful multi-parameter models were those stratified by car owners and non-car owners. Doubly-constrained versions of these stratified models performed better than the doubly-constrained single parameter versions.

These studies of the gravity model have shown that existing models do not provide reasonable estimates of base year trip tables. The errors in the base year are so large that it is recommended that the traditional trip distribution models not be used in urban transport planning studies. It is recommended that empirical studies of the determinants of commuting be continued and that the gravity model be used only to estimate interactions between stratified sets of origins and destinations. These strata would reflect socio-economic compatibility and timing of development compatibility in trip ends.

Acknowledgements

The data analyses on which this report is based were performed by Waterloo undergraduate students employed as work term computer analysts under the co-operative education program conducted by the University of Waterloo and a variety of employers. These students were W. Arnold (Mathematics), L. Brittain (Geography), C. Campbell (Mathematics), M. Dias (Systems Design), D. Hazzan (Systems Design), E. Hildebrand (Civil Engineering) and B. Underhill (Civil Engineering).

Several staff members from the Ontario Ministry of Transportation and Communications provided important advice throughout the preparation of the research report; these were V.C. Ma, P. Dalton and J. Saunders. L. Shallal of the Regional Municipality of Ottawa-Carleton and G. Thompson of the Regional Municipality of Waterloo kindly provided information on road network speeds for their respective municipalities.

The original manuscript was typed by Mrs. I. Steffler, Transport Group Secretary at the University of Waterloo.

References

[1] Regional Municipality of Waterloo, *Regional Transportation Plan Review: Study Design,* Report No. 2, August 1976.

[2] Bacharach, M., *Biproportional Matrices and Input Output Change,* Cambridge University Press, London, 1970.

[3] Furness, K.P., "Time Function Iteration," *Traffic Engineering and Control,* November 1965, pp. 458-460.

[4] Williams, W.T., and Lance, G.N., "Hierarchical Classificatory Methods," in Enslein, K., Ralston, A and Wilf, H.S., (Editors), *Statistical Methods for Digital Computers,* Wiley Interscience, New York, 1977.

[5] Baxter, R., and Williams, I., "An Automatically Calibrated Urban Model," *Environment and Planning,* Vol 7, No. 1, 1978, pp. 3–20.

Appendix A/ The Clustering Method

The clustering method used is based on Ward's method which uses the following Euclidean distance measure to compare the similarities in the destination vectors of residential census tracts:

$$d_{ij} = \left[\sum_{k=1}^{n} \frac{(T^*_{ik} - T^*_{jk})^2}{n} \right]^{1/2} \qquad \text{(A.1)}$$

where d_{ij} = the Euclidean distance between the destination characteristics

k = the destination zone under consideration for a particular origin zone (i or j)

T^*_{ijk} = the normalized number of trips between origin zone i and destination zone k

n = the total number of census tracts.

The clustering technique proceeds by first calculating an upper triangular matrix of the similarities in the destination vectors of each pair of census tracts. The pair of census tracts with the minimum d_{ij} magnitude are then merged into a cluster. The entries in the matrix of similarity are then recalculated to reflect this merging of two census tracts and the procedure is continued until all census tracts are merged into one cluster. The logic of the clustering procedure is illustrated in Figure A.1.

The calculations of the new entries in the similarity matrix do not involve a complete recalculation of the entries in the similarity matrix since not all entries are affected. Suppose two census tracts i and j join to form a new group ij, usually labelled with the smallest census tract number. The similarity between this new cluster and another cluster, or census tract, is given by:

$$d_{h,ij} = \alpha_i d_{hi} + \alpha_j d_{hj} + \beta d_{ij} \qquad \text{(A.2)}$$

where $\alpha_i = \dfrac{n_h + n_i}{n_h + n_i + n_j}$

$\alpha_j = \dfrac{n_h + n_j}{n_h + n_i + n_j}$

$\beta = \dfrac{-n_h}{n_h + n_i + n_j}$

n = the number of census tracts in the cluster identified by the sub-script.

Equation A.2 indicates that the new measure of similarity is calculated from the weighted sum of the Euclidean distances between the pairs of clusters used in the previous step. The larger n_i is relative to n_j then the greater d_{hi} will be weighted in the calculation of the new similarity measure $d_{h,ij}$. Further, the larger the value of n_h then the smaller the weighting of d_{ij} which is the distance between the two census tracts being merged.

The distance value calculated by Equation A.1 may be interpreted as the within cluster sum of squared differences between the destination vectors of two census tracts. As more census tracts are included in a cluster then the distance value as calculated by Equation A.2 represents the within-cluster sum of squared differences between a cluster centroid and the location of the individual census tracts in the cluster. The clustering procedure groups census tracts so as to minimize the increase in the sum of squared differences across the entire set of census tracts. One would expect that the first stage in the clustering hierarchy would consist of a number of pairwise census tract clusters each with a small increase in the sum of squared differences. Large increases in the aggregate sum of squared differences might be expected as clusters are grouped to form eventually a single cluster.

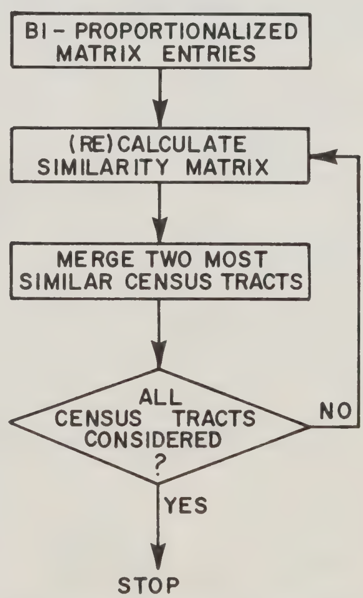


Figure A.1/ Steps in Clustering Procedure

Figure A.2(a) shows a simple five zone home to work matrix of the type available from the census. This matrix may be transformed into the doubly-balanced matrix shown in Figure A.2(b) by alternatively scaling the row and colum entries so that they sum to 20, an arbitrary constant.

The entries in the matrix of Figure A.2(b) may be used along with equation (A.1) to produce the upper triangular matrix of similarity measures shown in the upper part of Figure A.3. For example:

$$d_{12} = \left[\frac{(8-3)^2 + (7-8)^2 + (1-4)^2 + (4-4)^2 + (0-0)^2}{5} \right]^{1/2} = 2.645$$

and

$$d_{13} = \left[\frac{(8-4)^2 + (7-3)^2 + (1-7)^2 + (4-1)^2 + (0-7)^2}{5} \right]^{1/2} = 5.020$$

and so on.

Inspection of the initial similarity matrix of Figure A.3 shows that the minimum similarity coefficient is d_{35} with a magnitude of 2.49. The first step in the clustering procedure, then, is to group zones 3 and 5 to form a new zone 3′, where the marginal increase in the error be updated to reflect this initial clustering of zones 3 and 5 and this requires the calculation of $d_{13'}$, $d_{23'}$ and $d_{34'}$. Equation A.2 may be used to calculate these new similarities:

$$d_{13'} = \alpha_3 d_{13} + \alpha_5 d_{15} + \beta d_{35}$$

$$\alpha_3 = \frac{n_1 + n_3}{n_1 + n_3 + n_5} = \frac{2}{3}$$

$$\alpha_5 = \frac{n_1 + n_5}{n_1 + n_3 + n_5} = \frac{2}{3}$$

$$\beta = \frac{-n_1}{n_1 + n_3 + n_5} = -\frac{1}{3}$$

therefore

$$d_{13'} = \frac{2}{3} \times 5.02 + \frac{2}{3} \times 5.55 - \frac{1}{3} \times 2.49 = 6.22$$

$$d_{23'} = 5.54$$

$$d_{3'4} = 4.57$$

The revised similarity matrix which reflects the merging of zones 3 and 5 is shown in Figure A.3. It should be recalled that $d_{13'}$ represents the similarity between zone 1 and the new zone 3′ formed by the clustering of zones 3 and 5.

Inspection of the revised similarity matrix shows the minimum distance is $d_{12} = 2.65$ and the second step in the clustering procedure is to merge zones 1 and 2 to form a new zone 1′ with a marginal increase in error sum of squares of 2.65. A new similarity matrix has to be produced from the calculation of $d_{1'3'}$ and $d_{1'4'}$ and this matrix is shown in the lower part of Figure A.3. The minimum distance measure in this revised matrix is $d_{3'4}$ which means that the third clustering step involves merging 3′ and 4 to form a new zone 3″ with a marginal increase in the error sum of squares of 4.5. Once again a new similarity matrix has to be calculated which has one entry of $d_{1'3''} = 7.38$. The final fusion into one cluster occurs with a marginal increase in the error sum of squares of 7.38. The clustering sequence for this example is zones 3 and 5, then zones 1 and 2, then merged 3 and 5 with 4 and finally merged 1 and 2 with 3, 4 and 5.

Figure A.4 shows the dendrogram for this particular example where this dendrogram has been plotted using the convention of the computer program used in this study. The clusters are arranged as far as possible in terms of increasing zone number magnitude. The horizontal lines representing the fusion of two census tracts, or clusters, is drawn at the magnitude of the marginal increase in the sum of squares at which this union occurs where the magnitude is shown as the ordinate. That is, zones 3 and 5 are clustered first at a magnitude of 2.49; zones 1 and 2 are then clustered at a magnitude of 2.65, and so on. The lower part of Figure A.4 traces the marginal increase in the error sum of squares as the number of clusters decreases to 1. This diagram is simply a restatement of the vertical spacings of the horizontal lines shown in the upper part of the diagram.

		WORK ZONE					
		1	2	3	4	5	
HOME ZONE	1	14	3	1	5	0	23
	2	5	4	3	5	0	17
	3	10	2	5	1	2	20
	4	5	0	4	15	1	25
	5	6	1	2	4	2	15
		40	10	15	30	5	

(a)

(b)

		WORK ZONE					
		1	2	3	4	5	
HOME ZONE	1	8	7	1	4	0	20
	2	3	8	5	4	0	20
	3	4	3	6	1	6	20
	4	2	0	5	8	5	20
	5	3	2	3	3	9	20
		20	20	20	20	20	

Figure A.2/ Doubly-Balanced Home-to-Work Linkages Matrix

REVISED SIMILARITY MATRIX

HOME ZONE	1	2	3	4	5
1		2.65	5.02	5.02	5.55
2			4.31	4.27	5.25
3				4.05	2.49
4					4.07

INITIAL SIMILARITY MATRIX

HOME ZONE	1	2	3'	4	5
1		2.65	6.22	5.02	
2			5.54	4.27	
3'				4.57	
4					

REVISED SIMILARITY MATRIX

HOME ZONE	1	2	3'	4	5
1'			7.5	5.3	
2					
3'				4.5	
4					

Figure A.3/ Successive Similarity Matrices

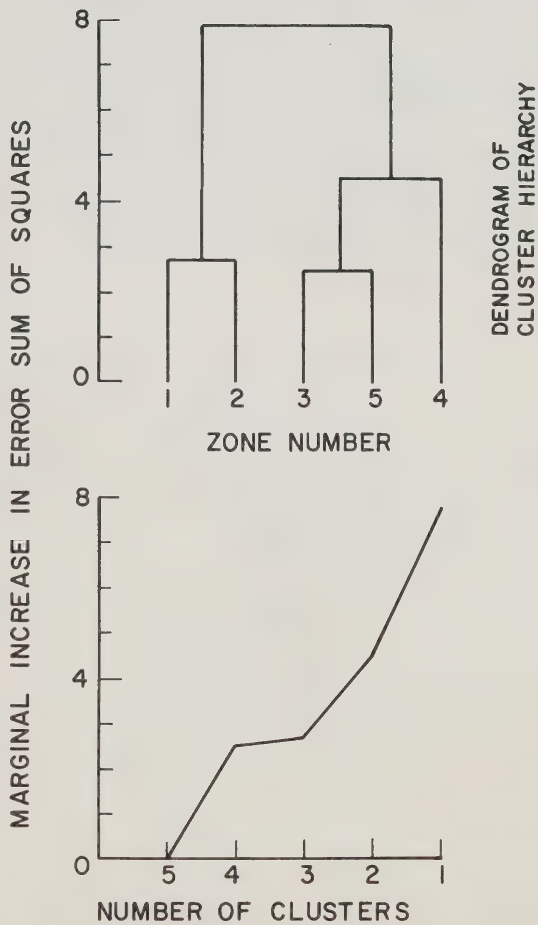


Figure A.4/ Dendrogram of Cluster Hierarchy and Associated Marginal Increase in Error Sum of Squares

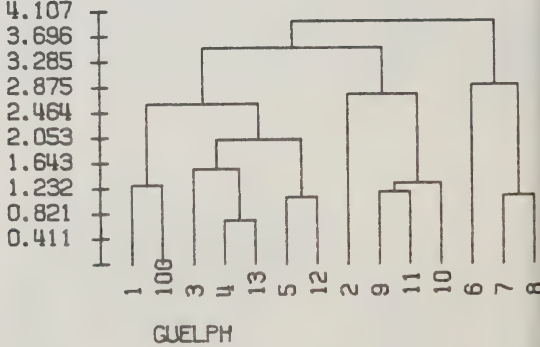
Appendix B/ Dendrograms and Cluster Boundaries for Census Areas

The dendrograms for each of the census areas are contained in this Appendix. In addition the boundaries of the major clusters or residential census tracts are also illustrated in this Appendix. The number of clusters ranges from three in Guelph to eight in the Ottawa CMA. The other piece of information contained in this Appendix are the major trip desires between home and work. Three diagrams are shown for each census area and the census areas are presented in order of increasing population.

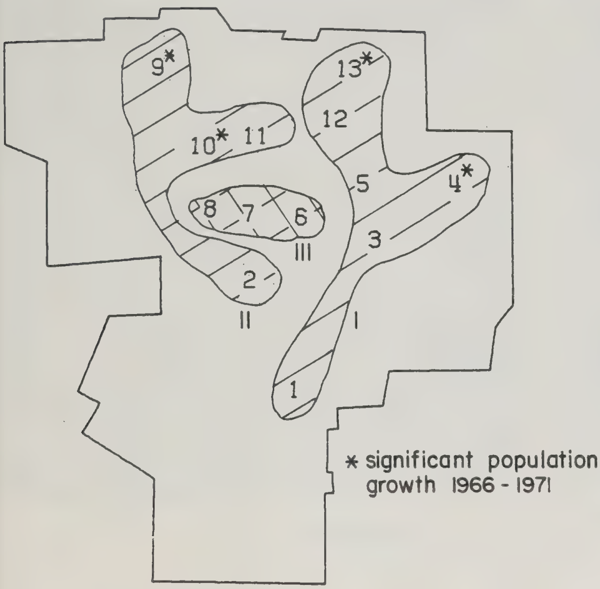
The dendrograms were plotted originally to the same scale in order that the rate of increase in the error sum of squares could be compared visually across all census areas. However, some of the dendrograms for the larger census areas had to be reduced in size in order to be incorporated in this report.

In the desire line diagrams only home to work linkage magnitudes of greater than 250 trips have been plotted in order to highlight the major trip desires. While this approach provides a reasonable view of the commuting patterns in the census areas with strong CBD concentrations of employment it tends to understate the commuting patterns in census areas such as Windsor and Kitchener where manufacturing employment is spread across a number of census areas.

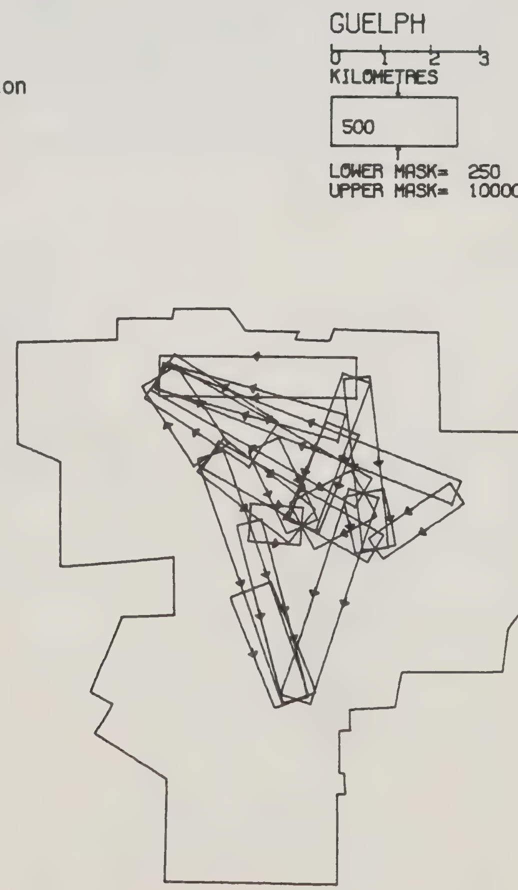
Guelph Census Area — The dendrogram for Guelph has the lowest increase in the error sum of squares of all census areas demonstrating that the destination composition of the residential census tracts is the most similar of all census areas. Three census tract clusters are identified for Guelph. Cluster I consists of the six census tracts along the eastern side of the area, cluster II consists of the four census tracts on the western side and cluster III embraces the CBD (CT 6) and the two census tracts immediately to the west of the CBD. The peripheral clusters I and II are linked with job opportunities in CT 1 (University of Guelph) and CT 9 (suburban industrial area) while the census tracts in cluster III are linked primarily to jobs in the CBD. The census tracts marked with an * experienced significant population growth during 1966–1971 and developed strong linkages with employment in CTs 1 and 9. The primary factor influencing spatial linkages in Guelph is the timing of development with the inner suburbs oriented to CBD employment and the newer outer suburbs linked with suburban employment. The major home to work linkages observed in Guelph in 1971 also illustrate these two broad commuting pattern differences.



Dendrogram for Guelph



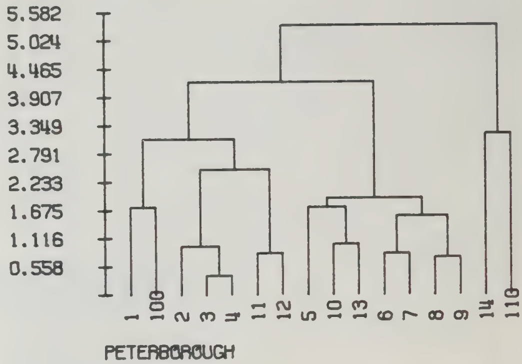
Census Tract Clusters for Guelph

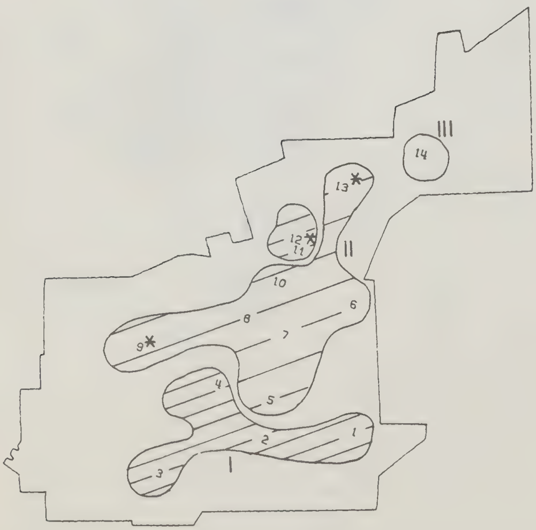


Major Home to Work Linkages for Guelph

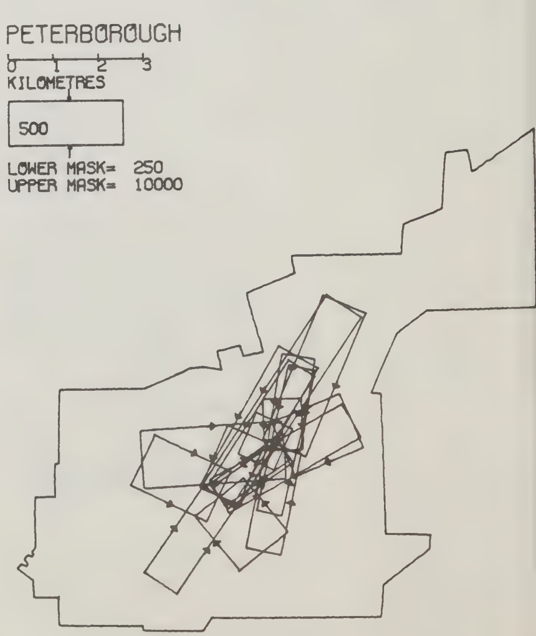
Peterborough Census Area — The Peterborough dendrogram shows that the destination composition of census tracts is fairly similar for all sections of the city. Three clusters can be identified for the urban area. CT 14 had a very small population in 1971 and it does not cluster naturally with the other Peterborough census tracts. The dendrogram shows that it clusters with CT 110 which includes the Lakefield Community north of Peterborough. Peterborough had two major concentrations of employment in 1971 and these were the CBD in CT 7 and the manufacturing employment in CT 4. Cluster I includes the older peripheral residential zones adjacent to the manufacturing plants in CT 4 along with CTs 11 and 12 which experienced significant population growth during 1966–1971. The principal determinants of this cluster are strong linkages with jobs in CT 4. In contrast the CTs in Cluster II have strong linkages with jobs in CT 7. The CTs in this cluster include the older established residential areas in CTs 5, 6, 7, 8 and 10 as well as the newer residential areas in CTs 9 and 13. The principal determinant of spatial linkages in Peterborough seems to be socio-economic with some influence of the timing of development.

Dendrogram for Peterborough





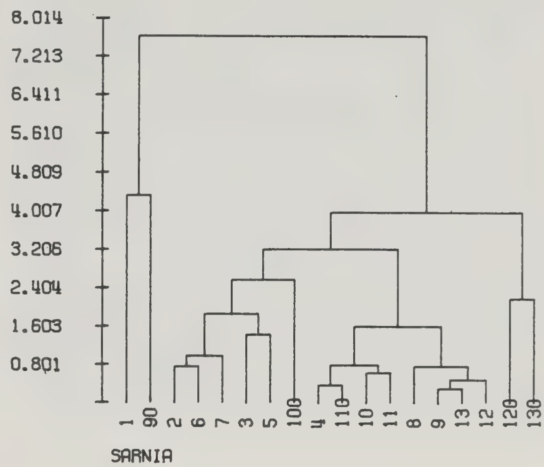
Census Tract Clusters for Peterborough

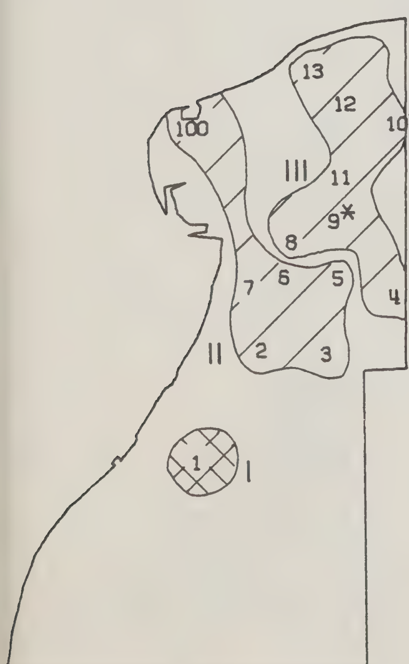


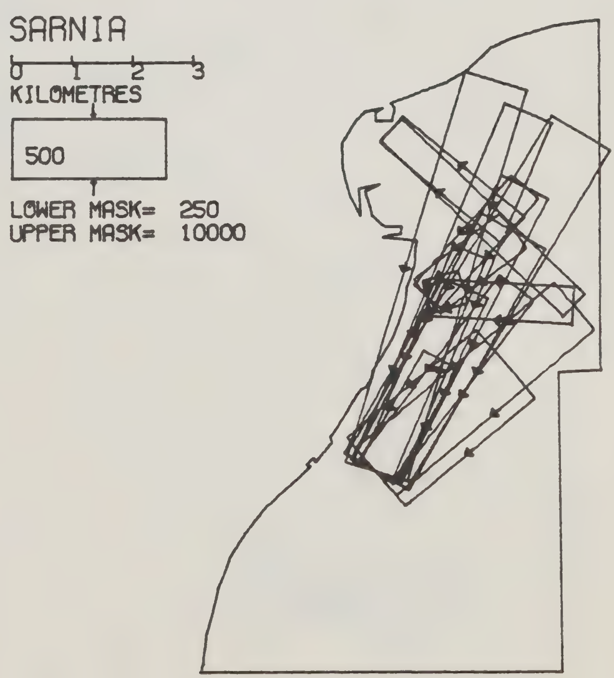
Major Home to Work Linkages for Peterborough

Sarnia Census Area — The dendrogram for Sarnia illustrates a significant increase in the error sum of squares with the addition of CT 1 and unofficial CT 90 to the rest of the city. Three major clusters can be identified. Cluster I consists only of CT 1 (and unofficial CT 90 not shown on the map) which contains the concentration of petro-chemical jobs in Sarnia and only a small number of households. Cluster II contains CTs 2, 3, 5, 6, 7 (CBD) and 100 (railway jobs). These CTs tend to have strong linkages with jobs in CTs 7 and 100. Cluster III embraces the remaining CTs along the eastern boundary of Sarnia which cluster on the basis of their strong linkages with jobs in CT 1. Some separation of the CTs in cluster III may be detected from the Sarnia dendrogram where differences are due primarily to linkages with jobs in CTs 7 and 100. The principal determinants of spatial linkages in Sarnia are socio-economic factors along with the strong concentrations of jobs in the petro-chemical industry. The major trip linkages in Sarnia highlight the importance of the three major employment concentrations.

Dendrogram for Sarnia





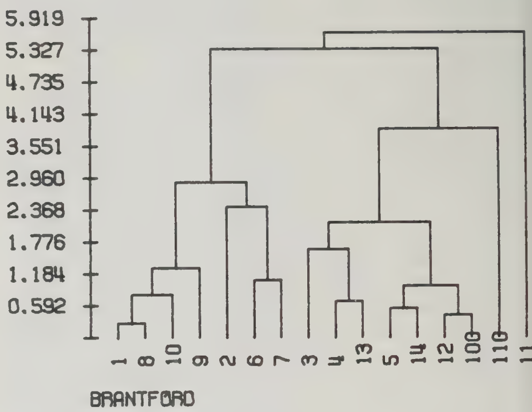


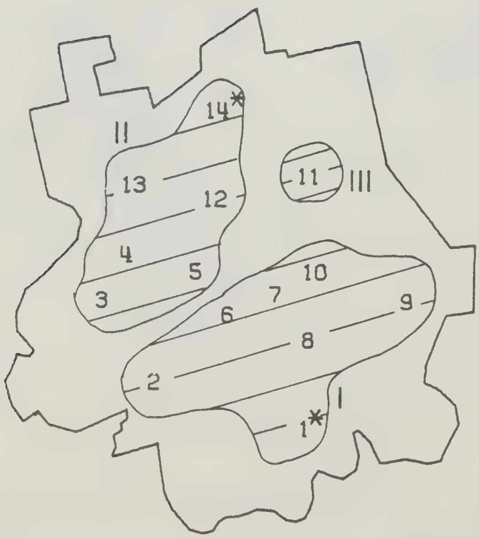
Census Tract Clusters for Sarnia

Major Home to Work Linkages for Sarnia

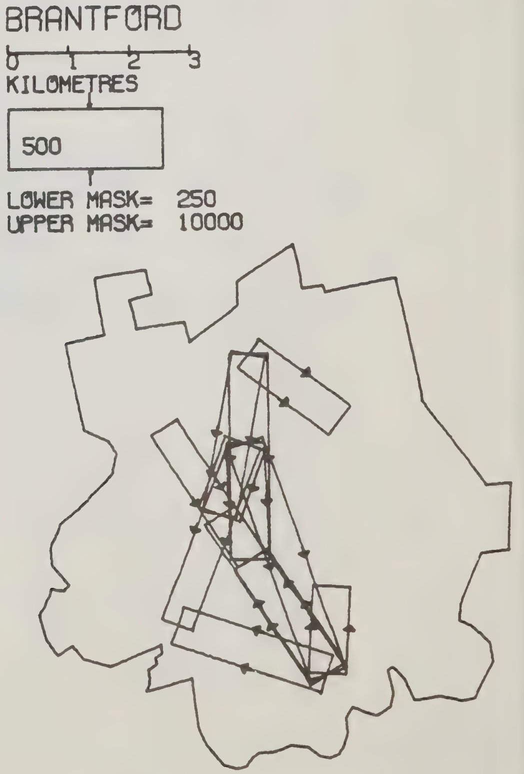
Brantford Census Area — The dendrogram for Brantford indicates that there are two distinct clusters of CTs which are joined only in the final stages of clustering. Three major clusters are identified where cluster III contains only CT 11, a major manufacturing zone. The cluster diagram indicates that the two major clusters are spatially separate from one another. While the significant amounts of employment in the central area, CTs 5 and 6, are common to both commutersheds the differentiation between the commutersheds is created by the linkages to jobs in CTs I and 2 for cluster I and linkages to jobs in CTs 3 and 11 for the CTs in cluster II. The dendrogram illustrates that there is a significant increase in the error sum of squares when clusters I and II join. This large increase in error sum of squares is due to CT 110 (Paris) joining cluster II just prior to its union with cluster I. The principal determinants of the clusters in Brantford are its multi-community composition and the presence of the Grand River which provides part of the boundary between clusters I and II.

Dendrogram for Brantford





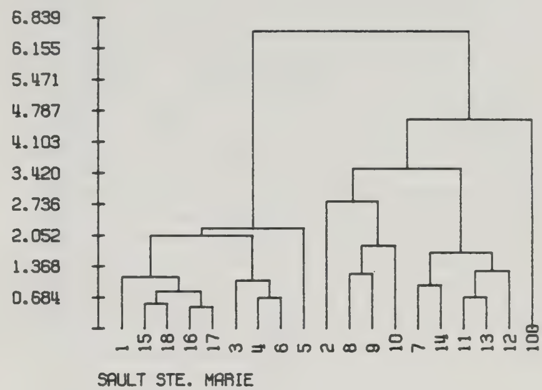
Census Tract Clusters for Brantford

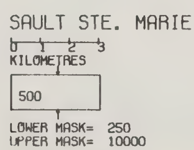


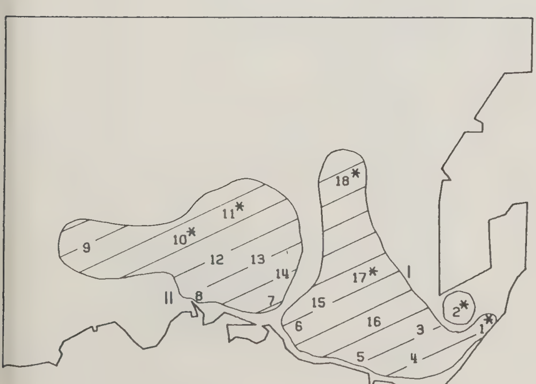
Major Home to Work Linkages for Brantford

Sault Ste. Marie Census Area — The dendrogram for this area shows two clearly identifiable clusters and their boundaries are outlined on the cluster map. Employment in 1971 was concentrated in CT 8 (steel mill complex) and CT 6 (CBD). Because of the magnitude of the jobs concentrated in CT 8 all residential zones have relatively strong linkages with this CT. The CTs in cluster I are in the eastern part of the area and they have strong linkages with the CBD employment as well as the steel mill complex. CTs in cluster II have very strong linkages to the steel mill complex. The dendrogram shows that there is more heterogeneity in the destination characteristics of CTs in cluster II than in cluster I and that there is a sharp increase in the error sum of squares when clusters I and II are merged. The principal determinant of commuting patterns in Sault Ste. Marie is the large concentration of jobs in the steel mill complex with socio-economic factors influencing some of the linkage patterns. The major trip linkages in Sault Ste. Marie show the dominance of the steel mill oriented trip linkages.

Dendrogram for Sault Ste. Marie





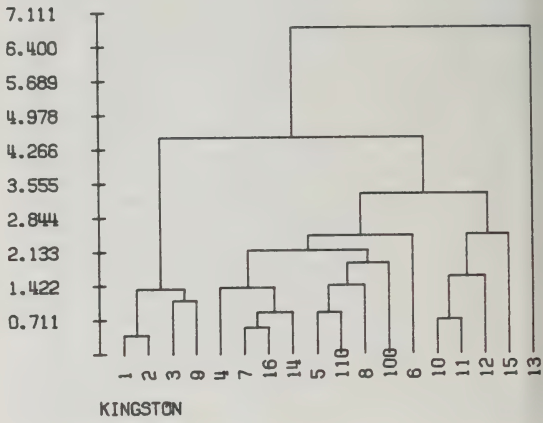


Census Tract Clusters for Sault Ste. Marie

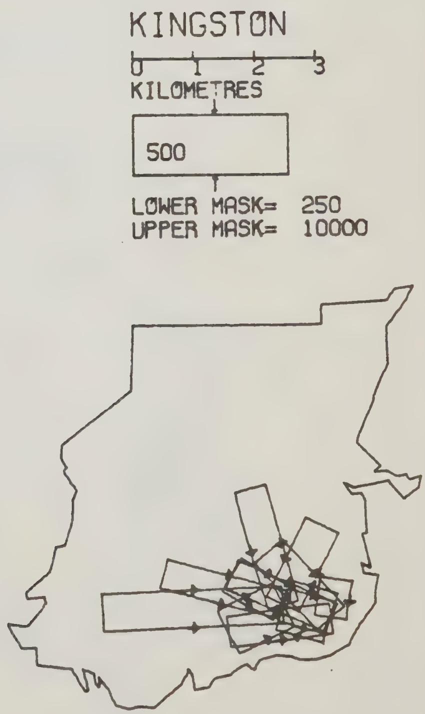
Major Home to Work Linkages for Sault Ste. Marie

Kingston Census Area — The dendrogram for Kingston shows that there is a sharp increase in the error sum of squares with the addition of CT 13. This CT is essentially a manufacturing employment zone which only clusters with the other CTs in the final stage. The cluster diagram shows three major clusters of census tracts for this area. Cluster I contains the CTs in the older inner suburbs of Kingston which have strong linkages with Queen's University (CT 2), the hospitals (CTs 2 and 1) and the CBD (CT 1). Clusters II and III embrace CTs that are located to the north and south of Princess Street. While the CTs in these two clusters focus on different employment centres there are strong socio-economic differences between the two clusters of census tracts. Socio-economic factors are the most important determinant of home to work linkages in Kingston. Topographic constraints also influence commuting linkages from the residential areas in Kingston Township but these are not shown. The major home to work linkages in Kingston illustrate the dominance of employment at Queen's University/Kingston General Hospital and the CBD.

Dendrogram for Kingston

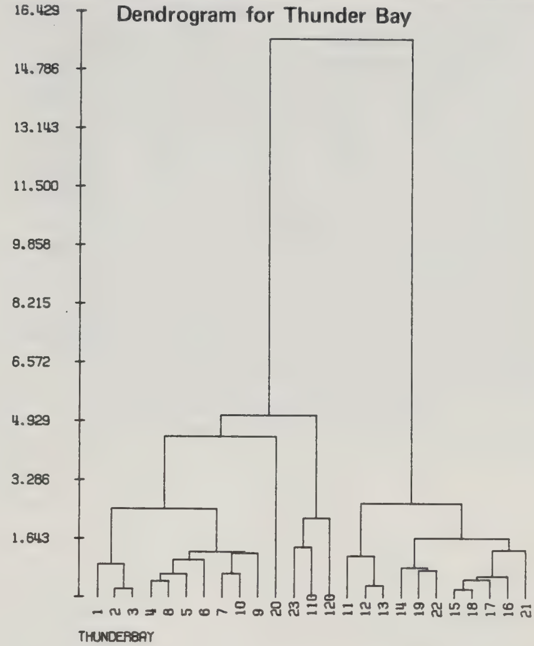


Census Tract Clusters for Kingston

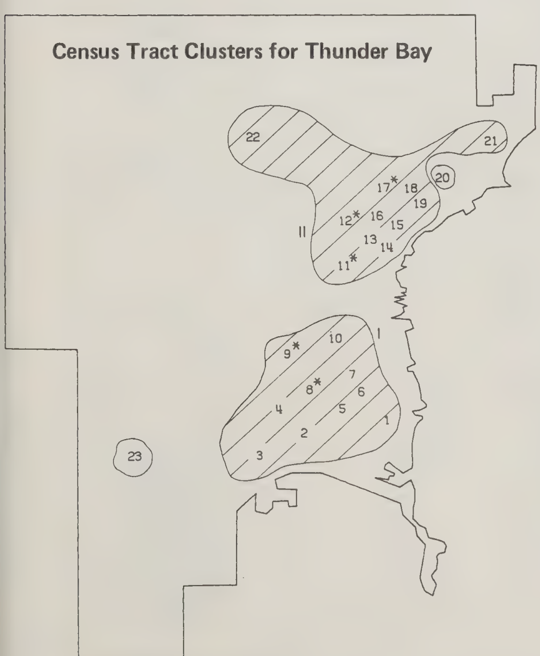


Major Home to Work Linkages for Kingston

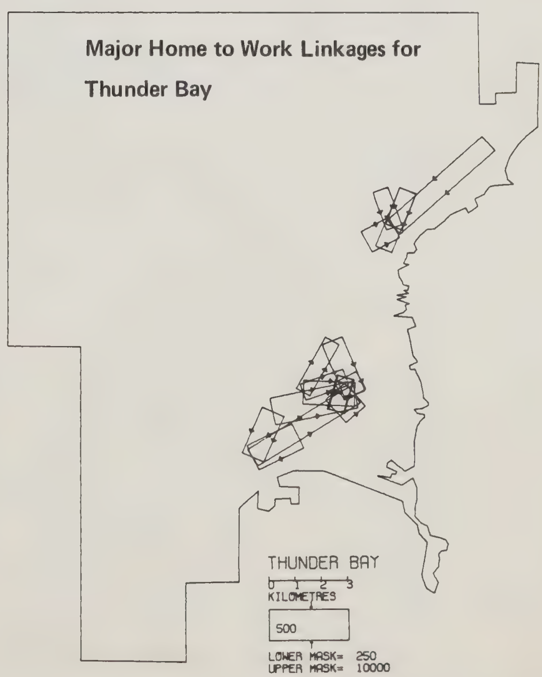
Thunder Bay Census Area — The dendrogram for Thunder Bay is the first in which there is a very large increase in the error sum of squares at any clustering phase. There are two major clusters of census tracts for Thunder Bay where these clusters embrace CTs within the former Cities of Fort William and Port Arthur, respectively. CTs 20 and 23 and outlying CTs 110 and 120 merge with the Fort William CTs in cluster I only after a significant increase in the error sum of squares. CT 20 is essentially an employment area and the residences in this CT exhibit an unusually long mean trip length. The dendrogram illustrates that destination patterns within each cluster are very homogeneous with the large increase in the error sum of squares occurring when clusters I and II are joined. Clearly, the major determinant of commuting linkages in the Thunder Bay area in 1971 were the former municipal boundaries. Within the former Fort William area employment in the CBD (CT 6) and CT 3 are the principal determinants of commuting patterns. The self-contained nature of the commuting patterns is further illustrated by the home to work linkage patterns.



Dendrogram for Thunder Bay



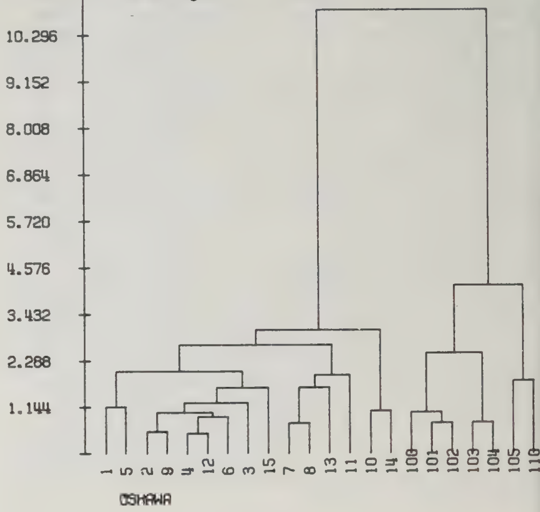
Census Tract Clusters for Thunder Bay



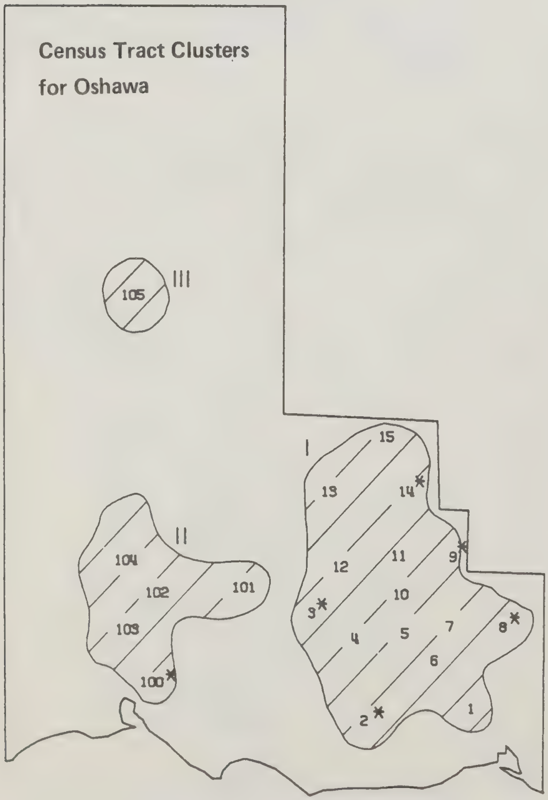
Major Home to Work Linkages for Thunder Bay

Oshawa Census Area — The dendrogram for Oshawa shows a major increase in error sum of squares at the last stage of clustering. The cluster diagram indicates that there are three major clusters in Oshawa where the boundaries of clusters I and II are essentially the municipal boundaries of Oshawa and Whitby, respectively. The dendrogram shows that cluster III (CTs 105, 110) merges with cluster II and that there is a sharp increase in the error sum of squares when clusters I and II join. Commuting patterns within Oshawa are dominated by the manufacturing employment concentrations. Municipal boundaries are the principal determinant of commuting patterns within the Oshawa census area and socio-economic factors tend to determine the spatial linkages patterns within Oshawa. In addition, the growth of housing opportunities in CT 105 north of Highway 401 has influenced the spatial linkages. The desire line diagram illustrates this and the municipal self-containment.

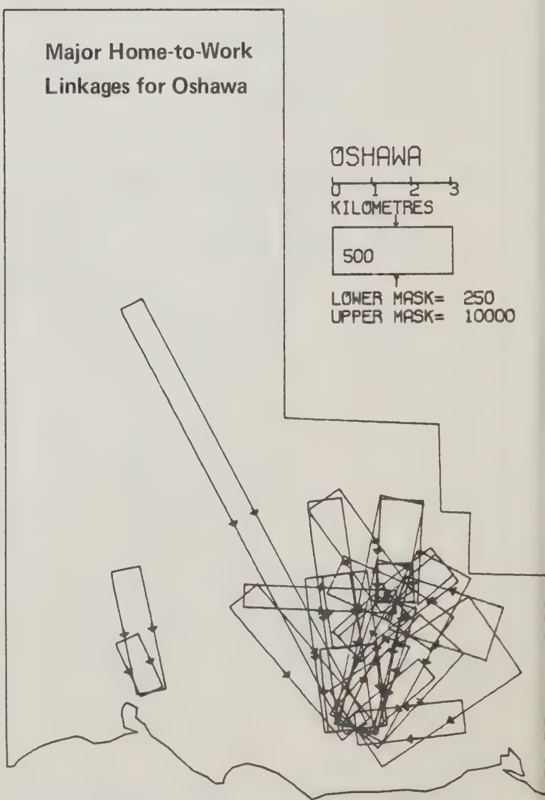
Dendrogram for Oshawa



Census Tract Clusters for Oshawa

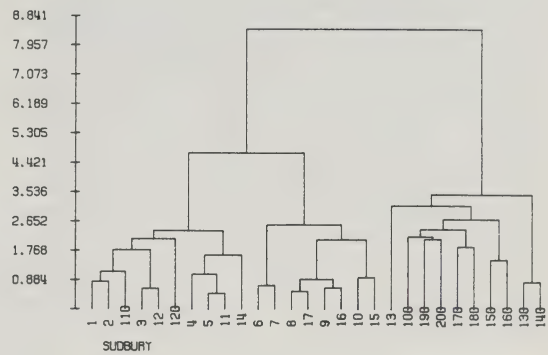


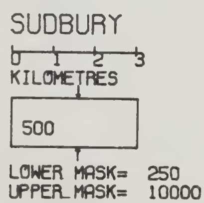
Major Home-to-Work Linkages for Oshawa

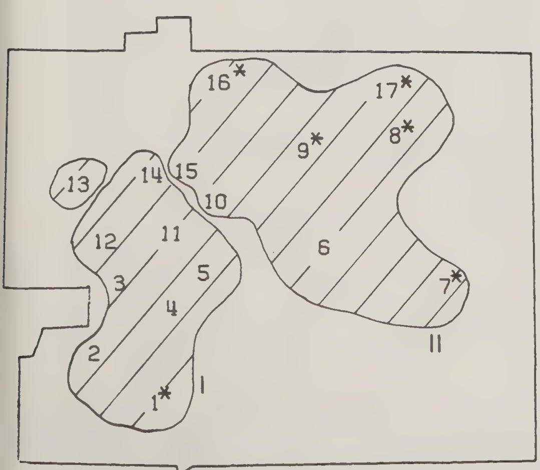


Sudbury Census Area — The Sudbury census tracts are grouped into two major clusters. The CTs in cluster I located in the southwest sector of the area have strong commuting linkages with the CBD. The CTs in cluster II are located in the northeast and tend to have strong linkages with the mining jobs. The dendrogram identifies a third cluster of residential areas which are located in the townships surrounding Sudbury. A significant increase in the error sum of squares occurs when clusters I and II are clustered with the CTs in the surrounding townships. It should be noted that the official census tracts in Sudbury cluster with a fairly low value of the error sum of squares indicating that the destination composition of census tracts within the city boundary is fairly homogeneous. The principle determinants of the commuting patterns in Sudbury are socio-economic differences and the locations of mining employment. The map of major trip desires indicated that the only significant linkages within the municipal boundaries are to the CBD. It should be noted, however, that there are major flows to the outlying mining centres.

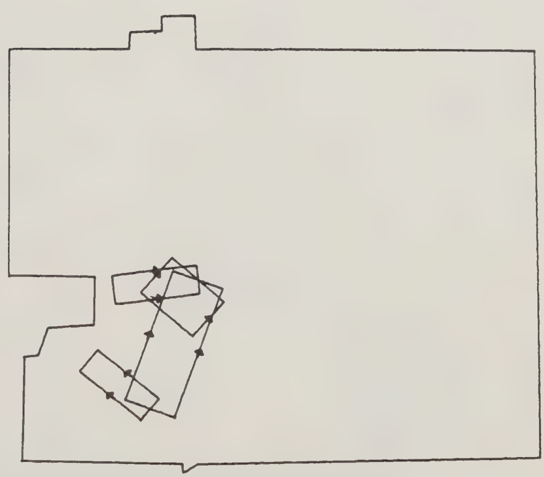
Dendrogram for Sudbury





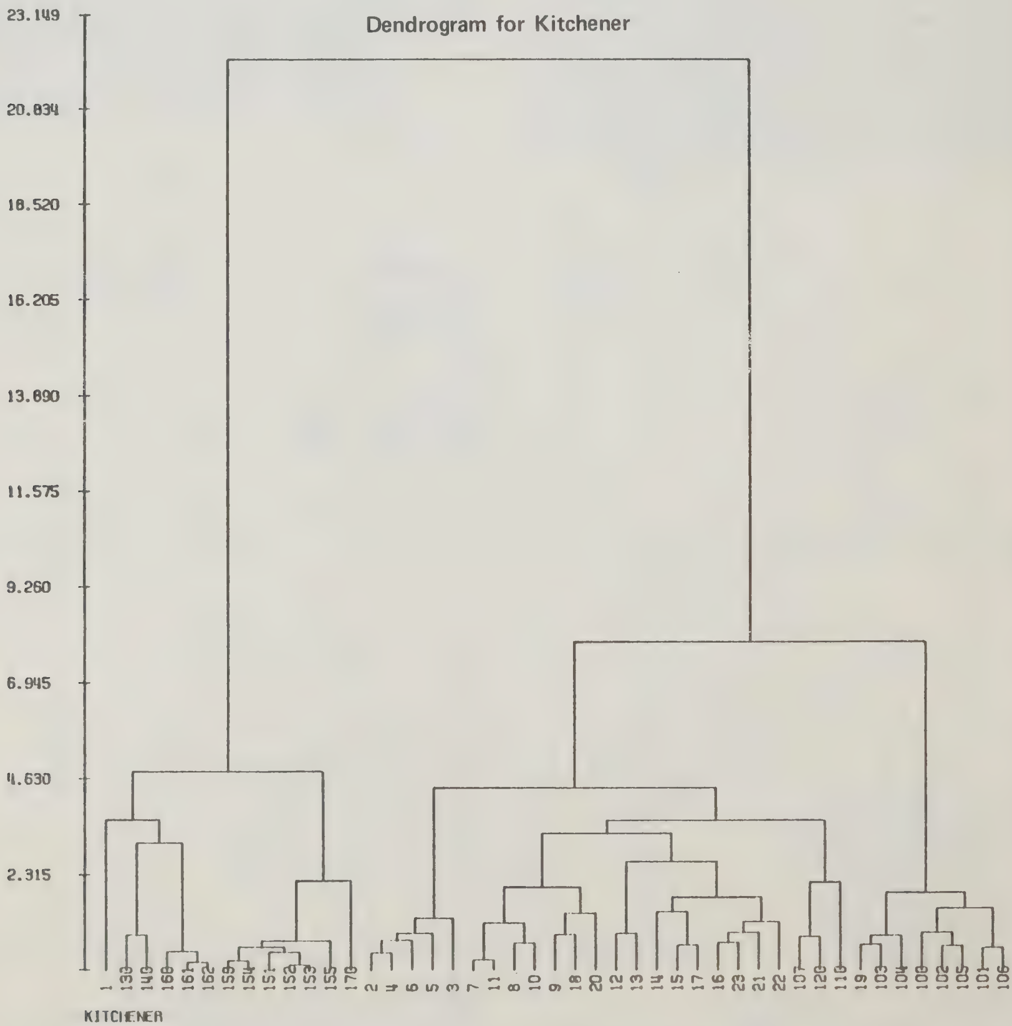


Census Tract Clusters for Sudbury

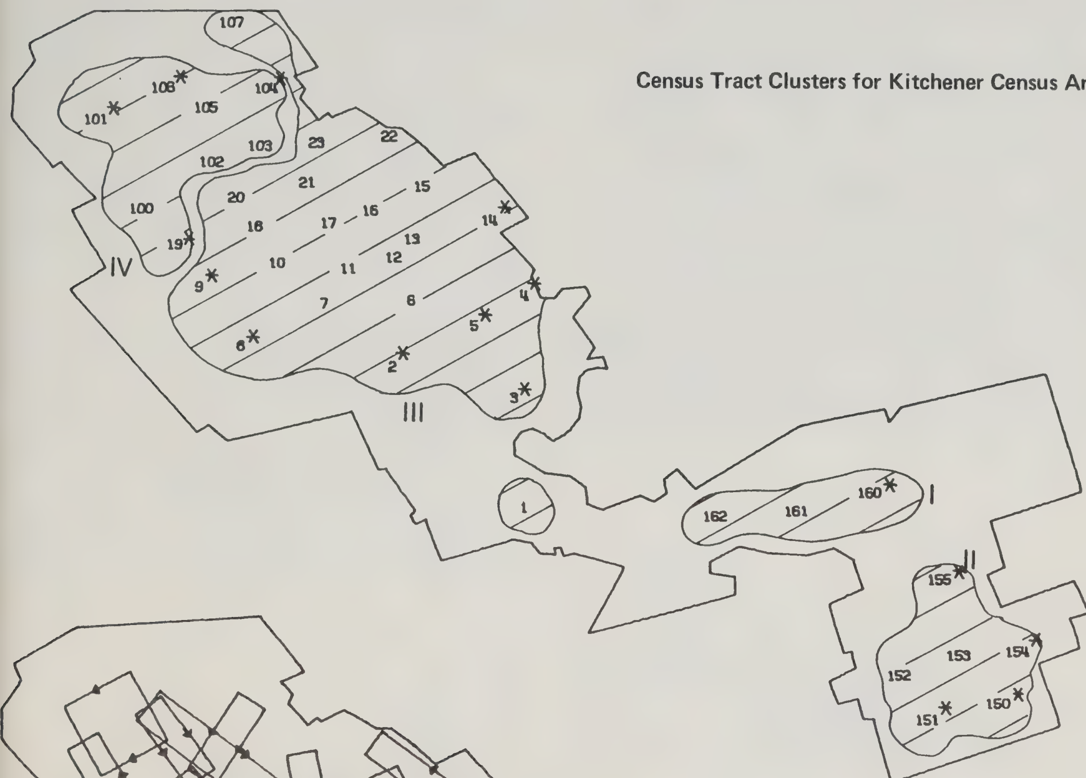


Major Home-to-Work Linkages for Sudbury

Kitchener Census Area — The Kitchener dendrogram clearly illustrates that there are major differences in commuting destinations between two large groups of census tracts. Four major clusters for the Kitchener area are illustrated on the cluster diagram where the cluster boundaries follow closely the former municipal boundaries of Waterloo, Kitchener, Preston and Galt. The dendrogram illustrates that the CTs in Preston and Galt cluster at about the same error sum of squares magnitude as the Kitchener CTs. A sharp increase in the error sum of squares occurs when the Kitchener and Waterloo census tracts are merged and the very high increase occurs when the Kitchener-Waterloo and Cambridge (Preston, Galt) clusters are merged. While manufacturing employment tended to dominate the economic base of the census area in 1971 there were major concentrations of non-manufacturing employment in the Kitchener CBD and at the University of Waterloo. As a result socio-economic factors have an important influence on commuting patterns in the census area as well as the municipal boundaries. The major trip desires in the Kitchener census area show fairly clearly the major commutershed sub-regions. Some of the linkages illustrated also reflect the timing of development.



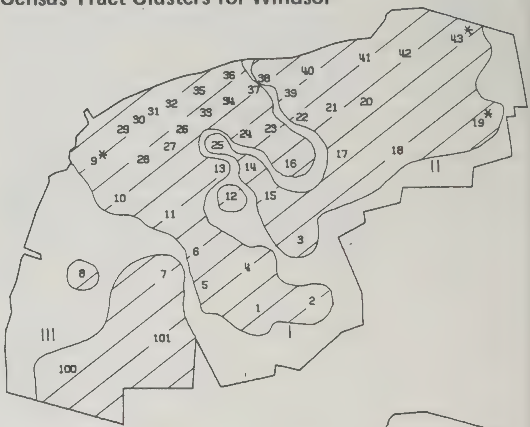
Census Tract Clusters for Kitchener Census Area

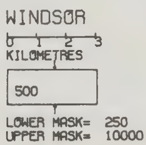


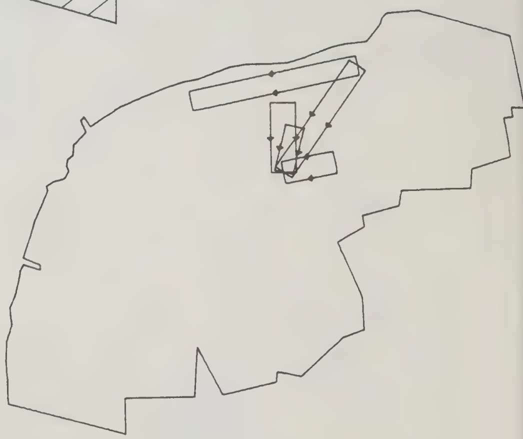
Major Home-to-Work Linkages for Kitchener Census Area

Windsor Census Area — The Windsor dendrogram indicates that the destination composition of residential census tracts is fairly homogeneous for a city this size. Three major clusters are identified on the cluster map. The CTs in cluster II to the east have strong linkages with CBD employment and with the automobile manufacturing plant (CT 16). The CTs within cluster I are more strongly linked to employment in CT 25. Cluster III consists of 13 CTs, only three of which are illustrated in the figure, with the remaining 10 being located in the townships outside of the city boundaries. The dendrogram illustrates that there is a significant increase in the error sum of squares when clusters I and II are merged and that there is a sharp increase when these are merged with cluster III. The principal determinants of trip linkages in Windsor are socio-economic factors, the timing of development and, to a lesser extent, the magnitude of employment in CT 16.

Census Tract Clusters for Windsor

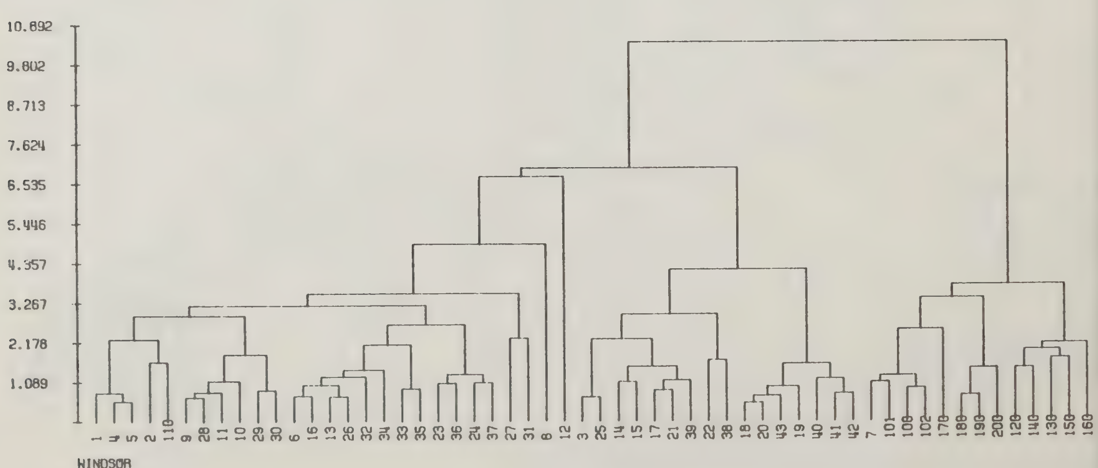




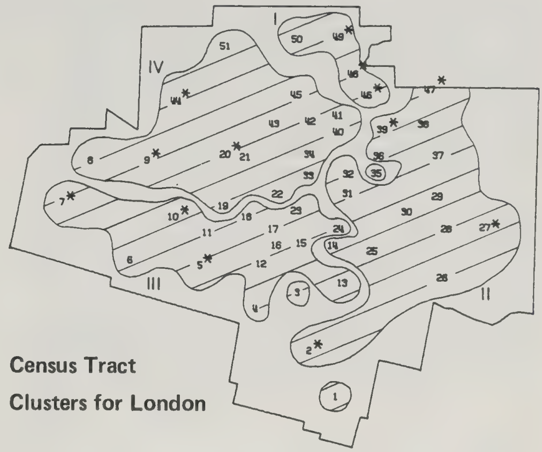


Major Home-to-Work Linkages for Windsor

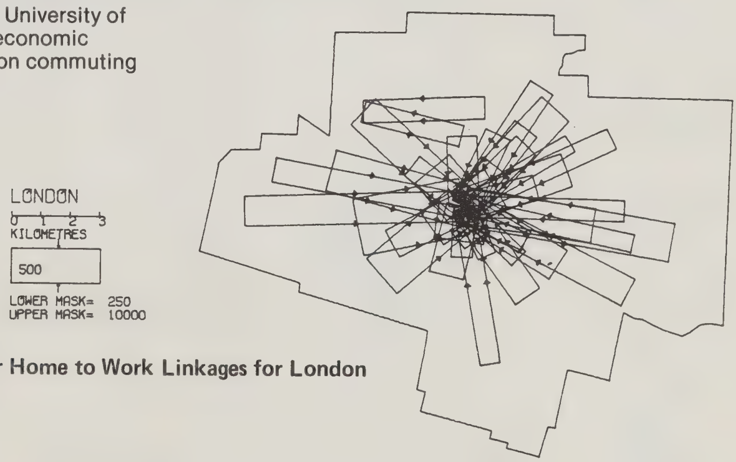
Dendrogram for Windsor



London Census Area — The London dendrogram illustrates that the city has a fairly homogeneous destination composition with a maximum value of the error sum of squares of about eight. Four major clusters for London are identified on the cluster map. Cluster I embraces the CTs on the northeast fringe of the area which are newer residential areas with commuting linkages to the CBD (CT 22). Cluster II consists of the CTs to the east of the CBD which also have strong linkages with the CBD and to employment opportunities in CT 27. Cluster III embraces CTs to the west of the CBD and south of the river. These CTs also have a strong orientation to jobs in the CBD. Cluster IV is to the north of the river and these CTs tend to have strong linkages with employment at the University of Western Ontario and to the CBD. The dendrogram shows that clusters I and II merge and clusters III and IV merge and that there is not a sharp increase in the error sum of squares when these merges occur. The major trip desires in London illustrate that commuting patterns are strongly dominated by the CBD employment and to a much lesser extent by the employment at the University of Western Ontario. The rivers and socio-economic factors also have an important impact on commuting patterns.

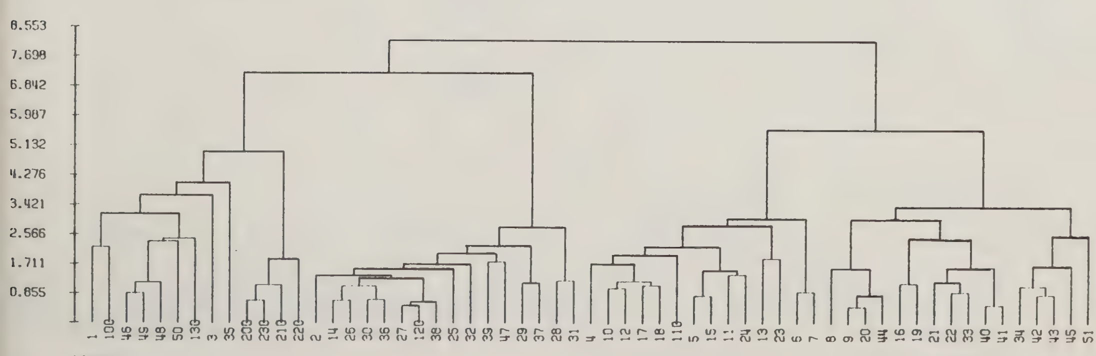


Census Tract Clusters for London

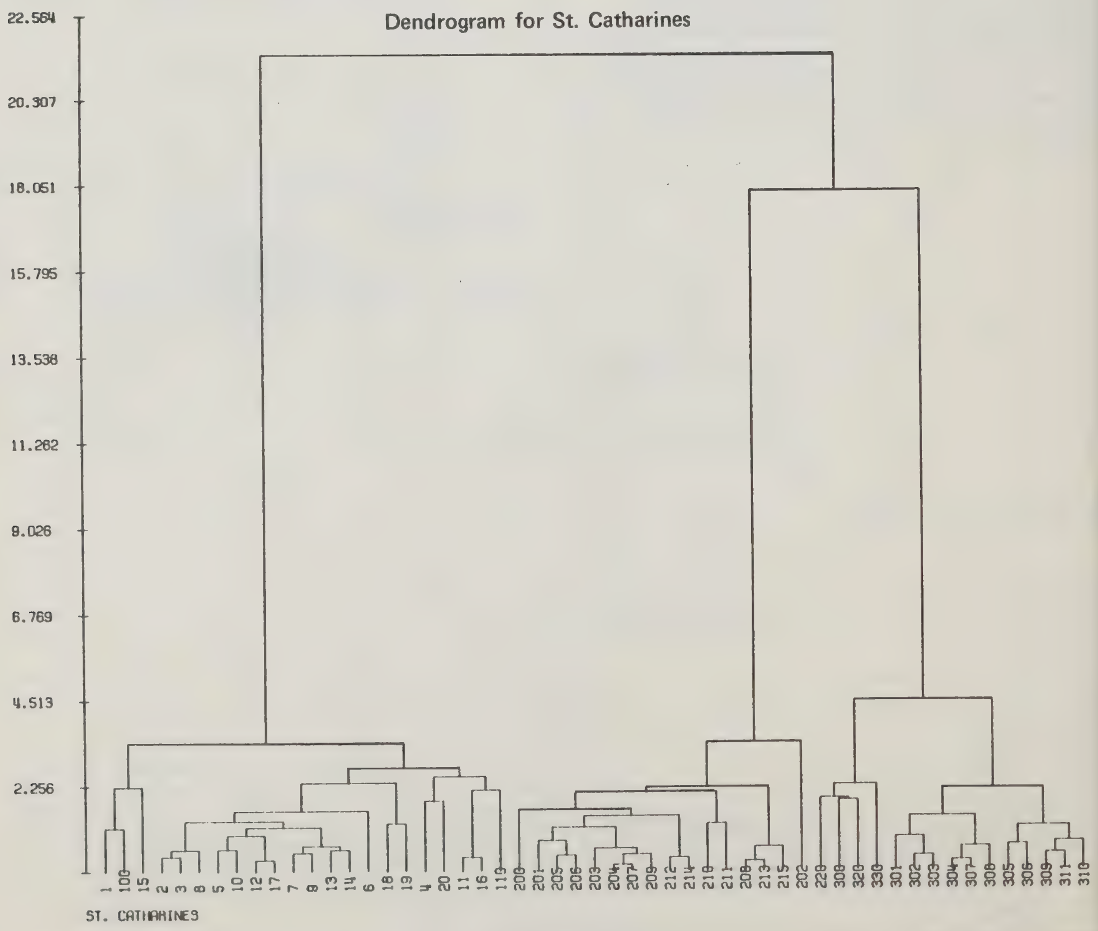


Major Home to Work Linkages for London

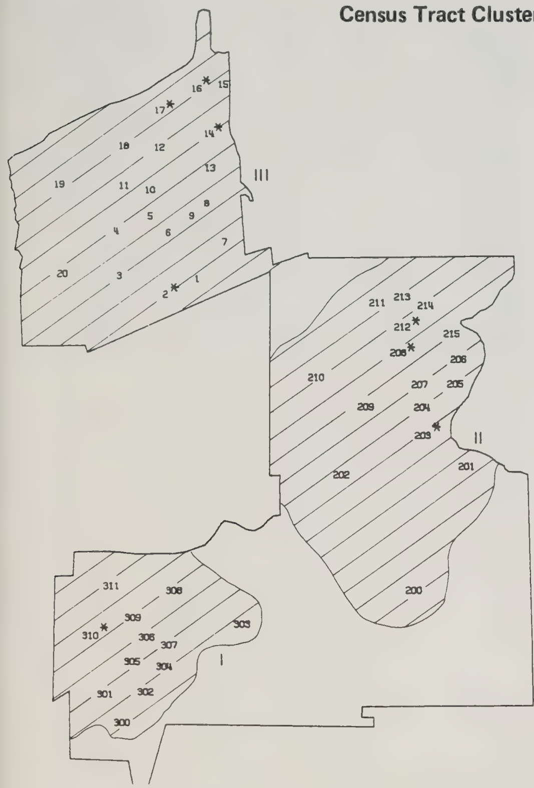
Dendrogram for London

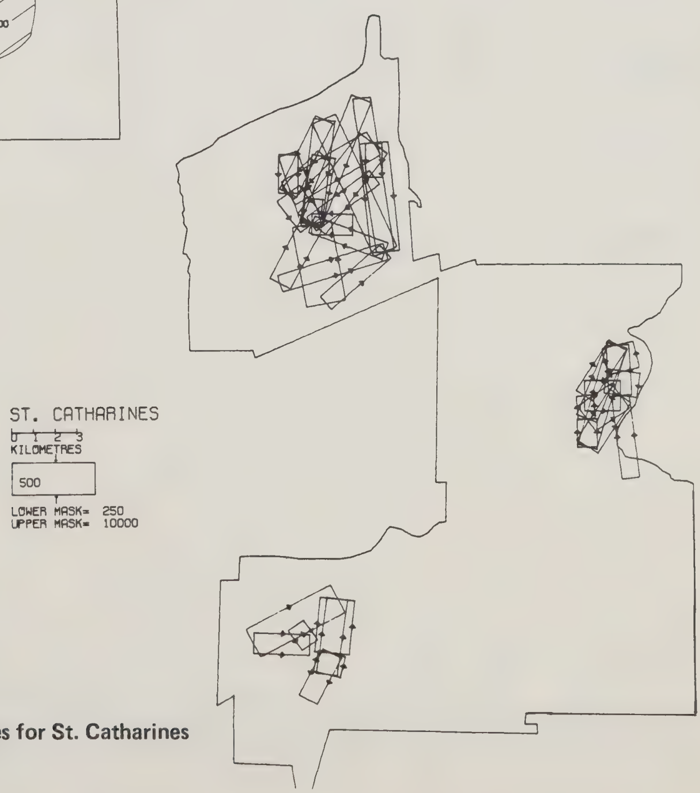


St. Catharines Census Area — The St. Catharines dendrogram illustrates that there are three major groupings of census tracts and that a very large increase in the error sum of squares is associated with the merges of these tracts. The cluster diagram illustrates the three major clusters of census tracts for the St. Catharines census area and this diagram shows that the cluster boundaries follow the municipal boundaries of St. Catharines, Niagara Falls and Welland. The dendrogram illustrates that it is the CTs within each municipality that cluster at low error sums of squares with the very large increase in the error sum when clusters are merged. Commuting patterns within each of the municipalities are determined principally by socio-economic factors. The trip desires for the area reflect the self-containment already suggested by the dendrogram.



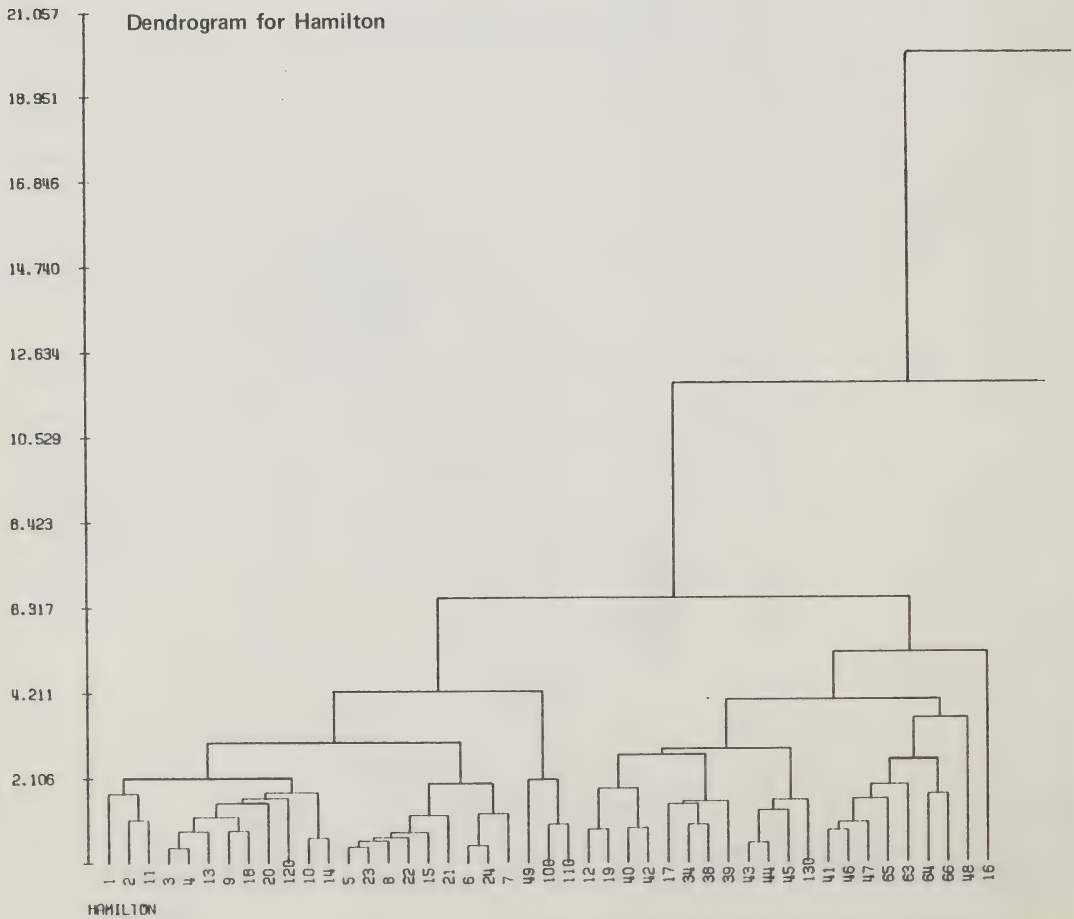
Census Tract Clusters for St. Catharines Census Area





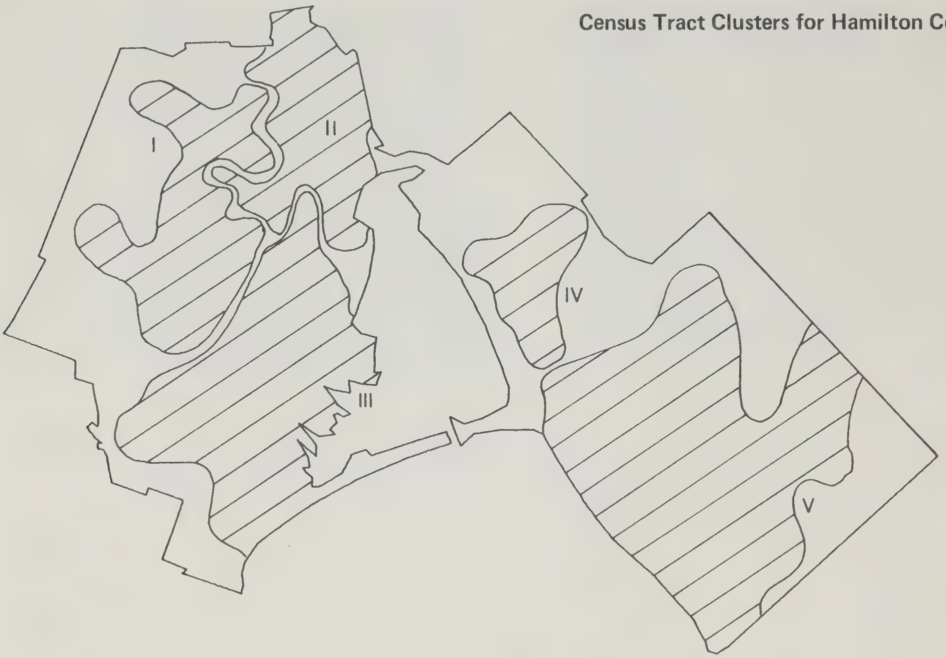
Major Home-to-Work Linkages for St. Catharines Census Area

Hamilton Census Area — The Hamilton dendrogram indicates that there are a large number of small clusters formed at a relatively low value of the error sum of squares. As these clusters join one another a number of large increases in error sum of squares occurs indicating the increasing heterogeneity of commuting destinations. The cluster diagram illustrates the census tract composition of five clusters for the Hamilton census area where three are located within Hamilton and two are within Burlington. The CTs in cluster I are located on the Niagara Escarpment and have strong linkages with the heavy industrial areas. Cluster II consists of CTs in the west end of Hamilton which have higher socio-economic status and these CTs have strong linkages with CBD employment and employment at McMaster University. Cluster II in the east end consists of lower status CTs which have strong linkages with the steel mill areas. The dendrogram illustrates that a significant increase in the error sum of squares occurs when clusters I and II are merged with cluster III and that very large increases occur when the Hamilton clusters are merged with the Burlington clusters. In Burlington the CTs in cluster IV have strong linkages with Hamilton while those in cluster V are more strongly linked with employment opportunities in Burlington. The desire line diagram illustrates the dominance of the industrial employment on commuting patterns. Job concentrations, topographic features and socio-economic factors are the primary determinants of commuting patterns in the Hamilton census area.



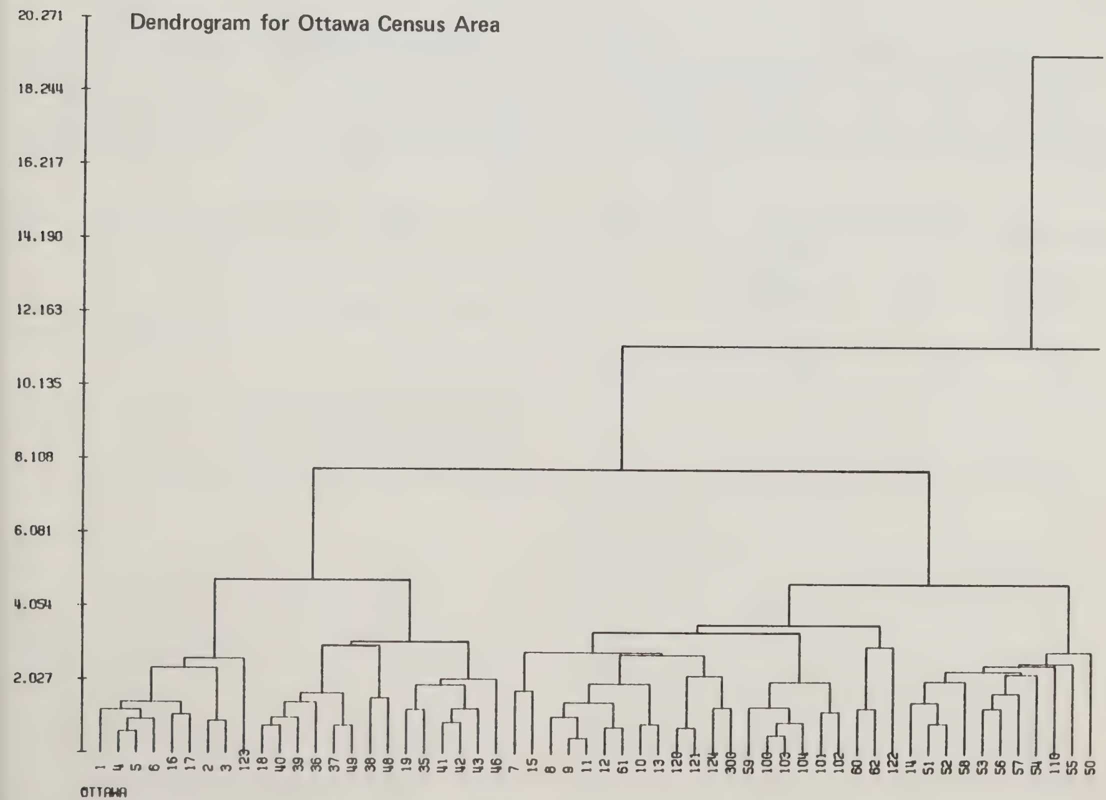

25 37 26 71 72 28 55 56 57 73 58 59 70 80 90 300 27 29 30 61 31 53 54 69 32 60 33 50 51 52 62 35 36 68 67 140 160 150 200 201 202 203 204 205 207 208 209 206 212 211 213 210 215 224 225 214 216 217 219 220 221 222 218 223

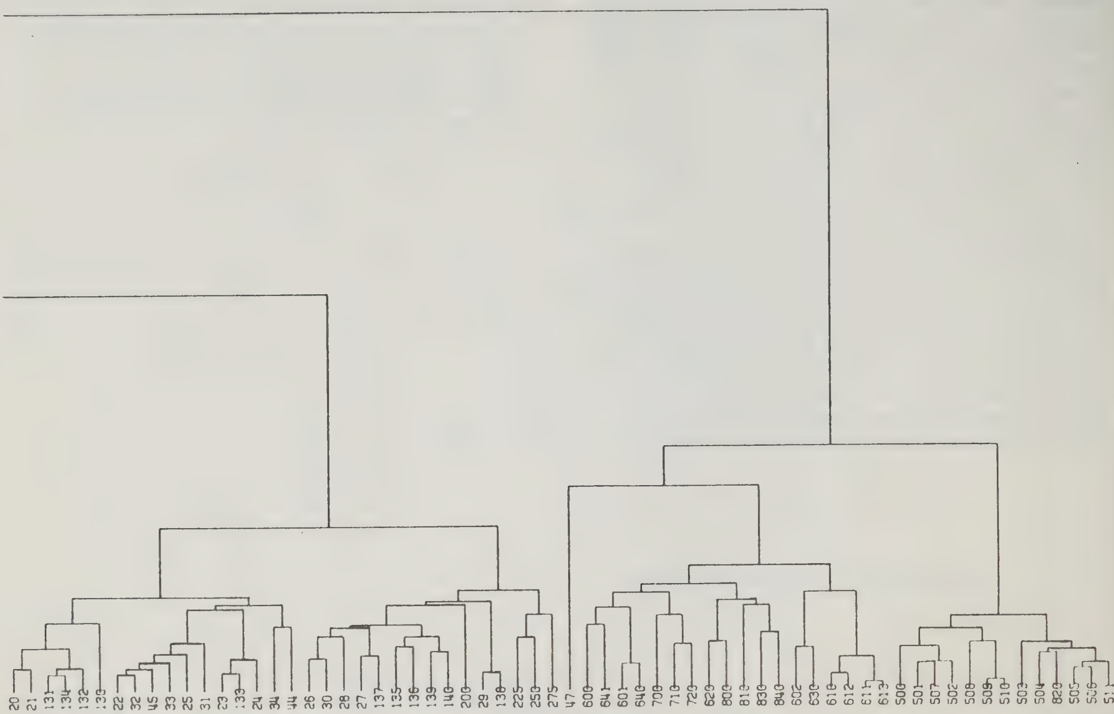
Census Tract Clusters for Hamilton Census Area



Major Home-to-Work Linkages for Hamilton Census Area

Ottawa Census Area — The Ottawa dendrogram has a pattern very similar to that of Hamilton where there are a large number of small fairly tight clusters which emerged with fairly regular increases in the error sum of squares. The cluster diagram illustrates eight census tract clusters for the Ottawa census area, six of which are within the Ottawa-Carleton region, and two of which are in Hull. Clusters II and IV embrace the older residential areas of Ottawa while the remaining six on the south side of the river consist primarily of the newer residential areas. The dendrogram shows that the clusters I and II which consist of CTs in a southerly corridor cluster, clusters III and IV in the east end cluster and clusters V and VI in the west end cluster. The next level of clustering illustrated by the dendrogram is the integration of clusters I, II, III and IV in the eastern half of the areas. There is a significant increase in the error sum of squares when all clusters on the south side of the river are joined and a sharp increase when clusters on both sides of the Ottawa River are joined. Commuting to the central area tends to dominate the commuting pattern as shown in the desire line diagram. In addition the timing of development has a strong influence as do socio-economic characteristics and topgraphic features.





Dendrogram for Ottawa Census Area

Appendix C/ The Gravity Model Program

C.1/ Model Formulation

Most of the analysis described in this report was done using a gravity model calibration program developed at the University of Waterloo. This program has the capability of calibrating a number of forms of the gravity model using a golden section search technique. The basic model utilized in the research is

$$T^*_{ij} = A_i B_j O_i D_j f(c_{ij}) \tag{C.1}$$

where T^*_{ij} = the simulated number of linkages between CT_i and CT_j

$$A_i = \left[\sum_j B_j D_j f(c_{ij})\right]^{-1} \tag{C.2}$$

$$B_j = \left[\sum_i A_i O_i f(c_{ij})\right]^{-1} \tag{C.3}$$

O_i = the employed labour force in CT_i

D_j = the employment in CT_j

$f(c_{ij})$ = a function of travel cost.

This is the doubly-constrained version of the gravity model. The terms A_i and B_j ensure that the restrictions

$$\sum_i T^*_{ij} = O_i \tag{C.4}$$

$$\sum_j T^*_{ij} = D_j \tag{C.5}$$

hold. A model with $A_i = 1.0$ is an attraction constrained model where only the restriction

$$\sum_i T^*_{ij} = D_j$$

holds. With $B_j = 1.0$ a production constrained model results and the restriction

$$\sum_j T^*_{ij} = O_i$$

holds. The calibration program can handle any one of these model types.

The travel cost function $f(c_{ij})$ can take one of two forms in the present model. The inverse power form of the model is

$$f(c_{ij}) = c_{ij}^{-\alpha}$$

where c_{ij} is the travel cost

α is a parameter to be estimated

The negative exponential function is of the form:

$$f(c_{ij}) = \exp^{-\beta c_{ij}}$$

where exp is the base of the natural logarithms

β is the parameter to be estimated.

The deterrence function may have a single parameter as described here or have separate parameters for calibration sub-regions. A multiple-parameter negative exponential function has the form

$$f(c_{ij}) = \exp^{-\beta_k c_{ij}}$$

where the index k identifies the calibration sub-region. It should be noted that these sub-regions are specific to origin zones; thus, trips produced in a particular zone i associated with sub-region k may be simulated using parameter β_k while those terminating there are simulated using the parameter corresponding to their origin zone.

The model used in a specific analysis is calibrated by minimizing some figure of merit which is a function of the differences between an observed and a simulated linkage matrix. Three measures are used and are listed below.

1. The absolute difference between global mean trip lengths.

$$\min \left| \frac{\sum_i \sum_j T^*_{ij} c_{ij}}{\sum_i \sum_j T^*_{ij}} - \frac{\sum_i \sum_j T_{ij} c_{ij}}{\sum_i \sum_j T_{ij}} \right| \tag{C.6}$$

2. The absolute difference between 1 km interval ordinates of the observed and simulated trip length frequency distribution.

$$\min \sum_{\ell=1}^{L} \left[\sum \sum \left[T^*_{ij} - T_{ij}\right]\delta_\ell\right] \tag{C.7}$$

where $\delta_\ell = 1$ if c_{ij} is in distance interval ℓ

$= 0$ otherwise

L = the number of 1 km distance intervals (max 50)

3. The sum of absolute difference between individual entries in the trip matrix

$$\min \sum_i \sum_j \left| T^*_{ij} - T_{ij} \right| \tag{C.8}$$

All work reported here utilized the measure of differences in the trip length frequency distributions using thirty 1-km intervals.

C.2/ Model Summary Output

The gravity model output has two basic purposes. The first is the identification of statistical and spatial regularities in the trip linkage patterns and the second is the analysis of model goodness of fit. Tables C.1 through C.6 contain example output for a single-parameter doubly-constrained model calibrated for home-owners in Thunder Bay. The model uses a negative exponential deterrence function. Table C.1 indicates the number of interchanges in which the observed and simulated values differ by a specified percentage or absolute value. This is very useful for

identifying outliers particularly when the absolute error is large. The % error is a less reliable indicator because many of the interchanges may be very small.

Table C.2 gives information on mean trip lengths by zone. The observed MTLs indicate the overall characteristics of trip lengths in an urban area and the simulated values make it possible to determine how well the gravity model simulates trip lengths in a spatial context. The expected mean trip lengths indicate the value of the trip length which would occur if the gravity model parameter were zero, the case of no distance deterrence. This indicates the general degree to which the model parameter affects the linkage pattern. It should be noted that these values are calculated both for originating and terminating trips at each zone.

Table C.3 lists the zonal components of the chi-square and phi-statistics calculated to assess model fit. The chi-squared statistic is calculated for the individual origin and destination zones as:

$$X^2_{origin} = \sum_i \frac{(T_{ij} - T^*_{ij})^2}{T^*_{ij}} \qquad (C.9)$$

$$X^2_{destination} = \sum_j \frac{(T_{ij} - T^*_{ij})^2}{T^*_{ij}} \qquad (C.10)$$

The corresponding calculation for phi is:

$$phi_{origin} = \sum_i T_{ij} \left| \log_e \frac{T^*_{ij}}{T_{ij}} \right| \qquad (C.11)$$

$$phi_{destination} = \sum_j T_{ij} \left| \log_e \frac{T^*_{ij}}{T_{ij}} \right| \qquad (C.12)$$

This table makes it possible to assess variations in model fit in different parts of the urban area.

Table C.4 presents the aggregate statistics calculated for the urban area and the model. These values are calculated as follows:

$$R^2 = 1 - \frac{\sum_i \sum_j (T_{ij} - T^*_{ij})^2}{\sum_i \sum_j (T_{ij} - T)^2} \qquad (C.13)$$

$$R^2_{adj} = 1 - \frac{\sum_i \sum_j (T_{ij} - T^*_{ij})^2}{\sum_i \sum_j (T_{ij} - \overline{\overline{T}}_{ij})^2} \qquad (C.14)$$

where $\overline{\overline{T}}_{ij} = \frac{O_i D_j}{T}$

$$X^2 = \sum_i \sum_j \frac{(T_{ij} - T^*_{ij})^2}{T^*_{ij}} \qquad (C.15)$$

$$phi = \sum_i \sum_j T_{ij} \left| \log_e \frac{T^*_{ij}}{T_{ij}} \right| \qquad (C.16)$$

Both X^2 and phi are segmented into overestimation, underestimation, and intrazonal error.

Three likelihood statistics may be calculated and these are:

$$L_{obs} = \sum_i \sum_j T_{ij} \log_e \frac{T_{ij}}{T} \qquad (C.17)$$

$$L_{\beta = 0} = \sum_i \sum_j T_{ij} \log_e \frac{\overline{\overline{T}}_{ij}}{T} \qquad (C.18)$$

$$L_{sim} = \sum_i \sum_j T_{ij} \log_e \frac{T^*_{ij}}{T} \qquad (C.19)$$

Equation C.17 calculates the highest likelihood magnitude that can exist for a given urban area since it operates on the observed T_{ij} magnitudes. Equation C.18 calculates the magnitude of the likelihood that would be given by a trip distribution model without interaction deterrence effects. Finally, Equation C.19 calculates the likelihood for a model with optimum β and this likelihood magnitude should fall between the two extremes.

A variety of other error statistics may be calculated and these include the following traditional measures.

Mean Error

$$MERR = \sum_i \sum_j (T_{ij} - T^*_{ij})/n^2 \qquad (C.20)$$

Mean Absolute Percentage Error

$$MABSPC = \sum_i \sum_j \left| \frac{T_{ij} - T^*_{ij}}{T_{ij}} \right| \times 100/n^2 \qquad (C.21)$$

Standard Deviation of Residuals

$$SDRES = \left[\sum_i \sum_j \left\{ \frac{(T_{ij} - T^*_{ij}) - MERR}{n^2 - 1} \right\}^2 \right]^{1/2} \qquad (C.22)$$

Total Absolute Error

$$TABSERR = \sum_i \sum_j \left| T_{ij} - T^*_{ij} \right| \qquad (C.23)$$

Mean Absolute Error

$$MABSERR = TABSERR/n^2 \qquad (C.24)$$

The final information contained in this table is the total number of linkages. The number of intrazonal linkages and the number of cells in the linkage matrix.

Table C.5 identifies the calibrated model parameters, specifically A_i, B_j and β_k or α_k. The total accessibility is the sum of the inverse of A_i.

Table C.6 contains a summary of parameter values by calibration sub-region. It is most useful when zones are grouped within a set of sub-regions.

Table C.7 contains a warning identifying linkages which have simulated values of zero with a non-zero observed value. These interchanges cannot be used in the calculation of many of the error statistics because of division by zero so this warning indicates whether the problem is serious or not.

Table C.8 summarizes the observed and simulated labour force and employment. This provides a check on the data and it also indicates whether the iterative procedure for calculating the A_i and B_j values is working adequately. If there are major differences between observed and simulated values in a doubly-constrained model this indicates that the limit on the number of balancing iterations in the program should be changed.

Table C.1/ Percentage and Absolute Error Distribution for Gravity Model

Thunder Bay Trips by Home Owners — C(M)A — 1 Parameter (s)

Error Distribution

% Error (SIM - OBS) / OBS * 100			Absolute Error SIM - OBS		
500 -	250	5	1000 -	500	0
250 -	225	2	500 -	450	0
225 -	200	0	450 -	400	0
200 -	175	4	400 -	350	0
175 -	150	5	350 -	300	0
150 -	125	5	300 -	250	0
125 -	100	5	250 -	200	2
100 -	75	12	200 -	150	6
75 -	50	17	150 -	100	14
50 -	25	37	100 -	50	70
25 -	0	53	50 -	0	90
0 -	-25	39	0 -	-50	18
-25 -	-50	19	-50 -	-100	7
-50 -	-75	7	-100 -	-150	3
-75 -	-100	0	-150 -	-200	0
-100 -	-125	0	-200 -	-250	0
-125 -	-150	0	-250 -	-300	0
-150 -	-175	0	-300 -	-350	0
-175 -	-200	0	-350 -	-400	0
-200 -	-225	0	-400 -	-450	0
-225 -	-250	0	-450 -	-500	0
-250 -	-500	0	-500 -	-1000	0

Table C.2/ Summary of Zonal Mean Trip Length Information

Thunder Bay Trips by Home Owners — C(M)A — 1 Parameter (s)

Zonal Mean Trip Length Statistics

CT No.	Origin OBS	Origin SIM	MTL Expected	Destination OBS	Destination SIM	MTL Expected
1	5.1	5.3	8.0	8.8	6.8	8.7
2	3.7	4.1	6.8	4.0	4.6	6.7
3	4.5	4.7	8.2	5.8	5.2	8.7
4	4.3	4.5	6.6	4.6	4.8	7.0
5	3.5	4.0	5.7	4.6	5.2	6.7
6	4.7	5.0	6.1	5.1	5.1	6.2
7	4.9	5.0	6.5	5.7	5.3	6.6
8	5.1	5.5	7.1	3.4	5.3	7.2
9	4.6	4.6	6.7	4.6	5.0	7.5
10	5.7	5.6	8.1	3.2	4.7	9.2
11	5.0	4.7	7.2	5.5	5.1	7.9
12	7.4	5.1	8.2	6.5	6.4	9.4
13	6.9	6.0	10.2	6.4	6.0	10.5
14	9.1	9.6	12.0	6.7	8.1	11.8
15	9.6	8.3	12.1	8.6	7.0	13.0
16	10.3	10.5	16.2	4.1	7.4	16.0
17	14.8	8.4	21.9	10.3	8.7	23.1

Table C.3/ Summary of Zonal Chi-Square and Phi Statistics

Thunder Bay C(M)A 1 Parameter (s)
Trips by Home Owners

Zonal Statistics

	CHI-Square		PHI	
CT No.	Origin	Destination	Origin	Destination
1	124.25	1106.01	239.27	972.03
2	290.63	484.45	741.67	657.38
3	113.84	579.15	304.59	551.84
4	255.26	149.58	747.36	330.75
5	219.28	626.98	416.14	1277.42
6	207.71	244.29	547.41	521.09
7	178.58	319.79	312.97	547.84
8	456.79	200.62	860.40	244.84
9	300.61	257.03	567.39	720.68
10	354.06	127.61	661.66	158.48
11	373.03	167.63	387.73	330.01
12	239.75	160.67	74.19	335.59
13	620.77	189.16	514.41	437.66
14	196.40	89.95	363.22	180.01
15	333.27	89.74	413.83	153.06
16	354.64	339.22	355.18	279.15
17	646.75	133.74	255.20	64.77

Table C.4/ Summary of Aggregate Model Statistics

Thunder Bay C(M)A 1 Parameter (s)
Trips by Home Owners

Aggregate Statistics

R square value is	0.91
ADJ R square is	0.81
CHI–square value is	5265.60
Underestimates CHI-square is	3860.46
Overestimates CHI-square is	1405.13
Intrazonal CHI-square is	836.94
Underestimates of intrazonal	825.31
Overestimates of intrazonal	11.63
PHI value is	7762.41
Underestimates PHI is	5300.25
Overestimates PHI is	2462.16
Intrazonal PHI is	1471.54
Underestimates of intrazonal	1385.67
Overestimates of intrazonal	85.87
Likelihood of observed matrix	-130945.81
Likelihood at BETA = 0	-140215.19
Likelihood at optimum BETA	-133787.31
Observed MTL is	5.38
Simulated MTL is	5.40
MTL at BETA = 0	7.94
Mean absolute percentage error	33.21
Mean error is	-0.01
Std. deviation of residuals is	46.38
Mean absolute error is	28.43
Total absolute error is	8215
Total linkage is	27420
Intrazonal linkages is	5160
Number of cells is	289

Table C.5/ Balancing Factor and Model Parameter Summary

Thunder Bay Trips by Home Owners		C(M)A	1 Parameter (s)
Balancing Factors			
CT No.	A I Times 10000	B J	BETA
1	1.1910	1.27951	0.18235
2	0.9509	0.84796	0.18235
3	1.1454	1.05684	0.18235
4	0.9825	0.91087	0.18235
5	0.8771	0.92653	0.18235
6	1.0029	0.87907	0.18235
7	1.0408	0.89856	0.18235
8	1.1355	0.85752	0.18235
9	1.0085	0.93549	0.18235
10	1.2344	0.84763	0.18235
11	1.0562	0.97152	0.18235
12	1.2058	1.23584	0.18235
13	1.5331	1.31814	0.18235
14	2.5671	1.74869	0.18235
15	2.2373	1.67370	0.18235
16	4.0005	2.05978	0.18235
17	5.6372	5.82026	0.18235

Table C.6/ Summary of Parameter Values by Calibration Sub-Region

Thunder Bay C(M)A 1 Parameter (s)
Trips by Home Owners

Parameters Values Summary

The Census Tracts Included in the Calibration Subregions and the Parameter Values are as Follows:

Region	CT Nos.							BETA S
1	1	2	3	4	5	6	7	0.18235
	8	9	10	11	12	13	14	
	15	16	17					

Table C.7/ Warning About Zero Linkage

Thunder Bay C(M)A 1 Parameter (s)

** Warning ** From CT No.	To CT No.	Simulated Value is Equal to Zero OBS Value
16	17	15.
17	15	15.

Table C.8/ Summary of Observed and Simulated Labour Force and Employment

Thunder Bay C(M)A

	Observed		Simulated	
CT	Labour Force	Employ-ment	Labour Force	Employ-ment
1	630	1635	630	1635
2	3135	1920	3135	1923
3	1710	3015	1711	3015
4	3240	1485	3241	1485
5	1545	4530	1545	4530
6	2640	2400	2640	2401
7	945	2100	944	2101
8	3150	420	3149	422
9	1740	4590	1740	4590
10	2835	285	2836	285
11	1005	1035	1003	1035
12	45	1095	45	1095
13	1665	1530	1667	1529
14	1065	450	1065	450
15	1095	465	1098	464
16	660	255	659	254
17	315	210	315	209

C.3/ Detailed Analysis of Evaluation Statistics

The following section of this appendix examines the behaviour of the goodness of fit statistics described above for selected Ontario census areas. The behaviour of the statistics is examined both within individual census areas and between census areas. It is not possible to approach this problem of the identification of the best goodness of fit statistics in the traditional statistical sense. The magnitudes of the statistics are influenced by the number of zones, the total number of trips and the characteristic probability distributions of the observed and estimated trip interchange magnitudes. A simulation approach has been taken in order to evaluate the performance of the alternative statistics.

C.3.1/ The Simulation Method — Eight Ontario census areas were selected for the detailed evaluation of model statistics; Table C.9 lists these areas along with the population, number of census tracts and total number of observed trips. The areas selected represent a range of population sizes and spatial structures.

Simulated trip interchange matrices were calculated by multiplying the observed trip interchange entries by a random percentage error where these errors were generated from a rectangular distribution with mean of zero and a specified range. Six percentage ranges were used and these were 10, 25, 50, 75, 100 and 150%. The random number generator GGU3 from the IMSL subroutine library was used and a simulated matrix calculated from the following equation:

$$T_{ij} = T^*_{ij} + \delta . T_{ij} \times RND \times FACT \qquad (C.25)$$

where δ = a random number taking values of +1 or −1

RND = a random number where $0 \leq RND \leq 1.0$

FACT = allowable percentage error divided by 100.

The same starting value or "seed" was used for each set of simulations for each city and the estimated trip interchange matrix was simulated three times.

The use of the same starting numbers made it possible to evaluate each statistic with respect to the variations in the allowable error only. The multiple simulations were used to ensure that these evaluations were robust and were not dependent upon a particular sequence of random numbers that happened to be generated. The simulated trip interchange matrices were then adjusted so that the row and column totals equalled the observed labour force and employment magnitudes. In this sense the simulated trip interchange matrices were similar to those produced by a doubly-constrained gravity model.

Tabel C.9/ Summary of Effectiveness of Goodness of Fit Statistics

	R^2	R^2_{adj}	χ^2	PHI	Likelihood	Standard Deviation of Residuals	Mean Absolute Error
Sensitivity to Error Independent of Error	No	No	No	Yes	No	Yes	No
Level of Sensitivity	Low	Medium	Very High	High	High	High	High
Similar Magnitudes For all Cities	Yes	Yes	No	No	No	No	No
Consistent With Other Measures	No	No	No	Yes	Yes	No	Yes

C.3.2/ Behaviour of Alternative Statistics – Figure C.1 shows the variation in the R^2 and R^2_{adj} magnitudes with allowable error percentage for the eight census areas. The tabulated magnitudes are presented in the appendix of this chapter. This diagram shows that in all cases in the R^2_{adj} statistic is lower than R^2 and more sensitive to error increase. It is interesting to note that even when the allowable percentage error is 100% the R^2 magnitudes are greater than about 0.7. The diagram also illustrates that R^2 is not a very sensitive statistic for comparing the observed and simulated trip matrices in the larger cities. This is also the case for the R^2_{adj} in the larger cities except for Ottawa, where it continues to decrease with increasing error percentage.

The star marked on each plotted R^2 line indicates the R^2 magnitudes obtained for the gravity models calibrated for each of the eight census areas. The locations of these points would suggest estimated trip matrices with errors in the range of 60–70% except for Windsor, Hamilton and Ottawa where error magnitudes of around 90% are indicated.

Figure C.2 shows the variation in the chi-squared statistics with percentage error for the eight census areas; the tabulated values are presented in the appendix. This diagram illustrates that the chi-squared magnitude increases sharply beyond an allowable percentage error magnitude of 75%. Beyond an error magnitude of 100% the increase in chi-squared tends to level off and this reflects the fact that negative trip interchange estimates are not allowed and the under-estimations are constrained to a maximum of 100%. The stars shown on the diagram identify the chi-squared magnitudes for the calibrated models and these correspond generally to error magnitudes of 75–100%.

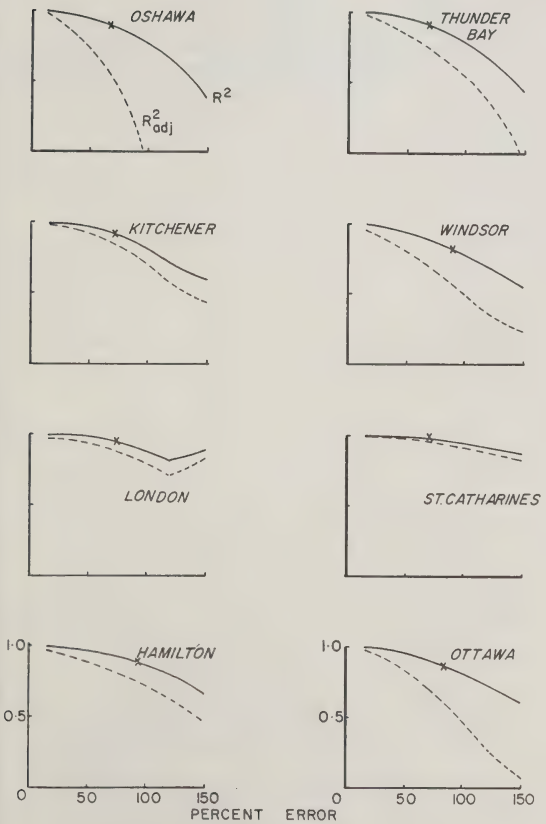


Figure C.1/Variation in R^2 and R^2_{adj} with Allowable Error Percentage

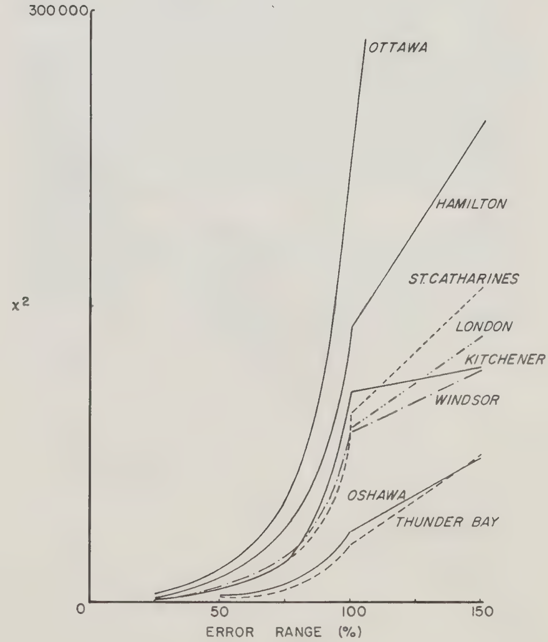


Figure C.2/Variation in Chi-Squared with Allowable Error Percentage

Figure C.3 shows the variation in the magnitude of the phi statistic with increasing percentage error. It is important to note that up to error magnitudes of 100% the phi statistic increases almost linearly with increasing percentage error. The phi statistic also appears to be more stable than chi-squared in that the ranking of cities is consistent across all error magnitudes except for Kitchener and Windsor. The phi magnitudes for the doubly-constrained models estimated for each census area are shown on the diagram suggesting interchange error magnitudes of the order of 80–90%.

Figure C.4 illustrates the variation in the likelihood statistics with increasing percentage error for the eight census areas; the tabulated magnitudes are presented in the appendix. The relationships illustrated show that the likelihood statistic behaves very similar to the chi-squared statistic with the magnitude percentages of about 75%.

Figure C.5 shows the standard deviation of the residuals with increasing error magnitude. This diagram illustrates that the increase is roughly linear with the exception of London where the magnitude decreases with a 150% allowable error. A similar behaviour may be observed in the R^2 statistic for London and this behaviour occurs only in the first of the three simulations and appears to be a result of the particular random number set generated. It is also interesting to note that the standard deviation decreases with increasing census area size: this is because the average size of an interchange is much larger for the smaller cities.

Figure C.6 shows the variation in the mean absolute error of a trip interchange magnitude with increasing trip interchange error. The mean absolute error behaves in a very similar way to the standard deviation of the errors.

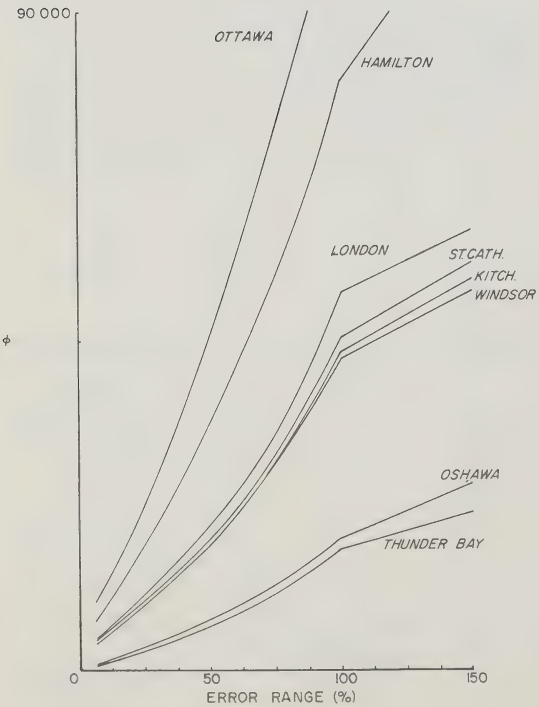


Figure C.3/Variation in Phi with Allowable Error Percentage

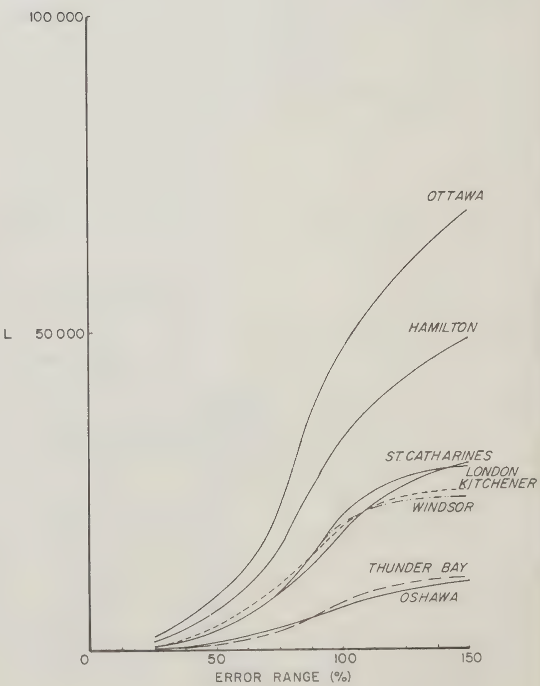


Figure C.4/Variation in Likelihood with Allowable Error Percentage

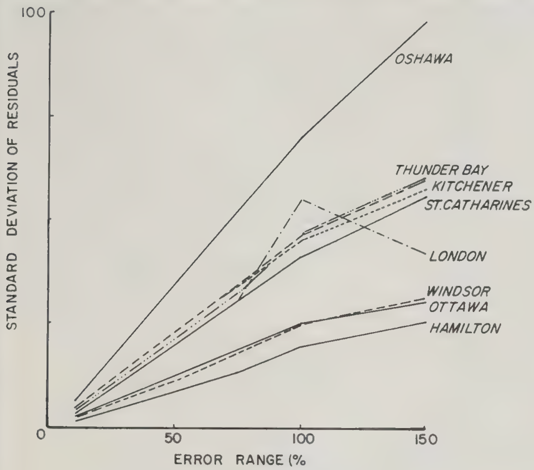


Figure C.5/Variation in Standard Deviation of Residuals with Allowable Error Percentage

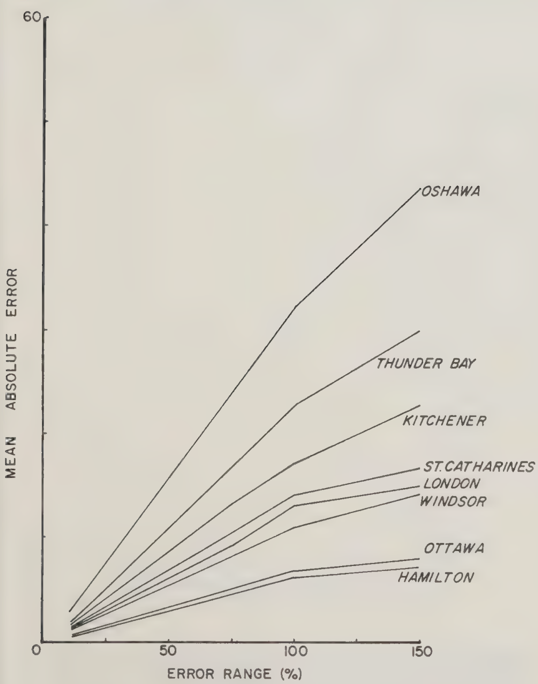


Figure C.6/Variation in Mean Absolute Error with Allowable Error Percentage

C.3.3/ Comparative Behaviour of Statistics – Three criteria may be identified for evaluating the quality of the alternative statistics and these are: (i) sensitivity to changes in error; (ii) similarity in magnitude across cities with different sizes; and (iii) consistency with other statistics.

The graphs of each of the statistics as a function of the maximum percentage error have shown that phi, the standard deviation of the residuals and the mean absolute error are equally sensitive to change in the error across all magnitudes of the error percentage. The chi-squared and likelihood statistics increase exponentially with increasing error magnitude and do not provide good discrimation at low error magnitudes. The R^2 statistic had very poor sensitivity to error magnitude changes remaining at high levels for large changes in the error magnitudes. The R^2_{adj} statistic is a more sensitive statistic for some of the urban areas but was insensitive in some urban areas particularly London and St. Catharines.

With the exception of the R^2 and R^2_{adj} statistics none of the statistics may be used to compare the goodness of fit of models across all urban areas since the magnitudes of the statistics are sensitive to urban area size, or the number of trip interchanges entries being compared. However, some transformations of the statistics are suggested in the following sections to allow comparisons across urban areas.

The third criterion suggested above is consistency. The chi-squared statistics appears to be inconsistent for large error magnitudes since the error functions cross-over at large percentage error magnitudes. Also, in one of the simulations, the R^2 and standard deviation of the residual measures are inconsistent for London where both magnitudes decreased when the error percentage increased from 100 to 150%.

Table C.9 summarizes the ratings of each of the statistics with respect to the three criteria. This would suggest that the phi statistic and the mean absolute error statistics are perhaps the best statistics to use. However, their principal deficiency is that they could be used across urban areas to compare model behaviour.

C.3.4 Comparisons Across Urban Areas — Chi-squared and phi statistics have been calculated for all of the Ontario census areas as well as for some stratified matrices for Kitchener, Hamilton and Ottawa using a maximum random error range of 75%. Figures C.7 and C.8 show the magnitudes of chi-squared and phi plotted against the number of home to work linkages in each area. These relationships indicate underlying linear trends suggesting that a linear transformation of either measure may be used as an indication of model performance across urban areas. Perhaps the most useful measure, then, is the phi statistic because of its constant sensitivity to error magnitude.

Figure C.9 shows R^2 and the phi statistic divided by the number of interchanges plotted against the number of interchanges when the estimated matrix has been generated using a 75% error magnitude. This diagram shows that the transformed phi statistic is quite consistent across all urban area sizes while there is a band of variation of R^2.

The behaviour of these two statistics for the calibrated gravity models is shown in Figure C.10. This graph shows that the transformed phi statistic increases with urban area size indicating that model performance is poorer with increasing urban area size.

Another possibility would be to use the standard deviation of the errors and the mean absolute error statistics to characterize the goodness of fit of trip distribution models. However, an evaluation of these statistics across urban areas indicated that they had certain undesirable characteristics. The analyses presented above would suggest that the transformed phi statistic is the best goodness of fit measure to use. This statistic is used only as an evaluative statistic in this report but its properties would suggest that it should be used as a calibration criterion as well.

C.4 Model Graphical Output — The model program will also produce graphical output if requested. Figure C.11 displays the observed and simulated trip length frequency distributions for a production constrained inverse power model for the City of Brantford. This model simulated linkages for individuals living in the same dwelling for six or more years. A second type of plot is the desire line representation of residuals. Figure C.12 illustrates the residual pattern for a doubly-constrained negative exponential model for Brantford. A third plot is shown in Figure C.13. The histograms plotted on the map indicate the production and attraction totals for home renters and home owners in Brantford. The capability exists to plot a variety of measures in this manner including parameter values, trip lengths, and balancing factors.

The calibration program is available on magnetic tape in both source and loaded form from the University of Waterloo. A users' manual is also available.

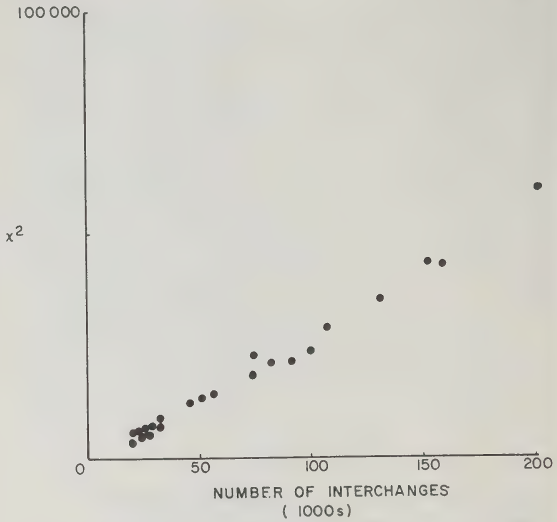


Figure C.7/Variation in Chi-Squared with Number of Trip Linkages for 75% Error Magnitude

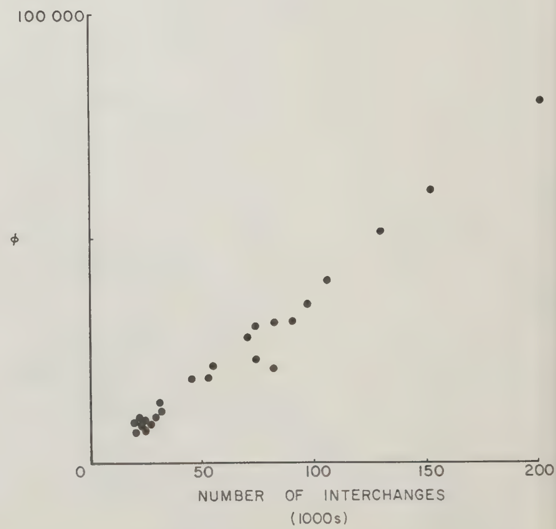


Figure C.8/Variation in Phi with number of Trip Linkages for 75% Error Magnitude

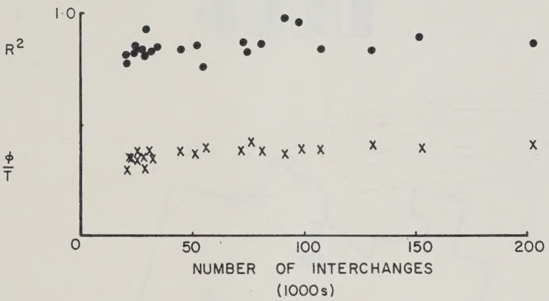


Figure C.9/Variation in R^2 and Phi/T with Number of Trip Interchanges for a 75% Error

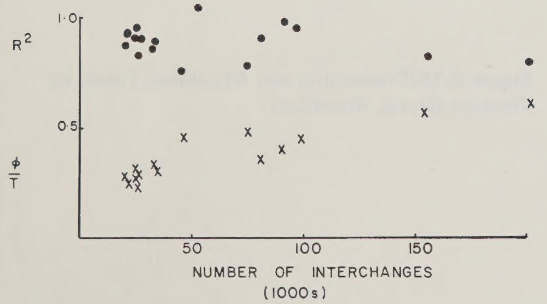


Figure C.10/Variation in R^2 and Phi/T with Number of Trip Interchanges for Single Parameter Doubly-Constrained Gravity Model

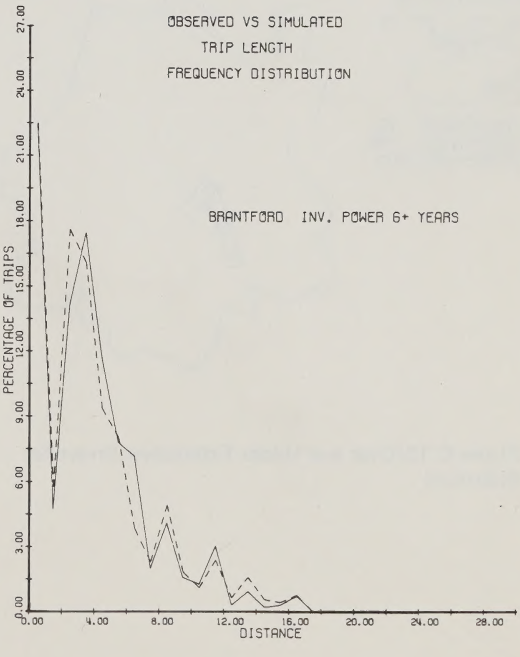


Figure C.11/Observed and Simulated Trip Length Frequency Distributions

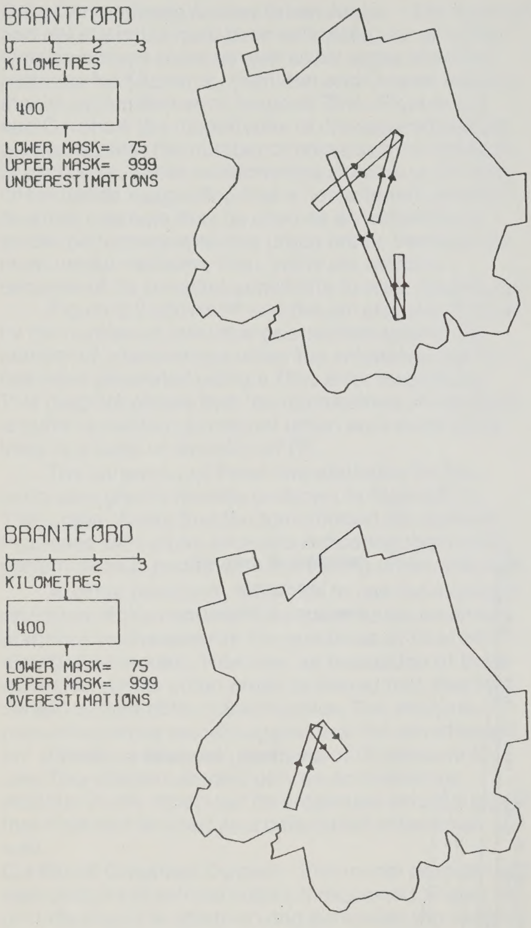


Figure C.12/Over and Under-Estimation Errors for Brantford

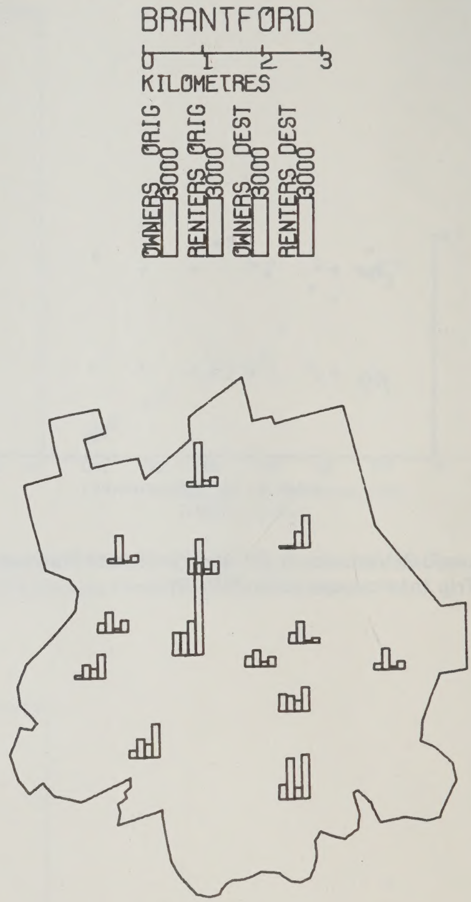


Figure C.13/Production and Attraction Totals by Tenancy Group, Brantford

RR 223

ISBN 0-7743-7565-5